SUR LES LIGNES GÉODÉSIQUES DES SURFACES CONVEXES*

PAR

HENRI POINCARÉ

§1. *Introduction.*

Dans mes *Méthodes nouvelles de la Mécanique céleste* j'ai étudié les particularités des solutions du problème des trois corps et en particulier des solutions périodiques et asymptotiques. Il suffit de se reporter à ce que j'ai écrit à ce sujet pour comprendre l'extrême complexité de ce problème; à côté de la difficulté principale, de celle qui tient au fond même des choses, il y a une foule de difficultés secondaires qui viennent compliquer encore la tâche du chercheur. Il y aurait donc intérêt à étudier d'abord un problème où on rencontrerait cette difficulté principale, mais où on serait affranchi de toutes les difficultés secondaires. Ce problème est tout trouvé, c'est celui des lignes géodésiques d'une surface; c'est encore un problème de dynamique, de sorte que la difficulté principale subsiste; mais c'est le plus simple de tous les problèmes de dynamique; d'abord il n'y a que deux degrés de liberté, et puis si l'on prend une surface sans point singulier, on n'a rien de comparable avec la difficulté que l'on rencontre dans les problèmes de dynamique aux points où la vitesse est nulle; dans le problème des lignes géodésiques en effet, la vitesse est constante et peut être regardée comme une des données de la question.

M. HADAMARD l'a bien compris, et c'est ce qui l'a déterminé à étudier les lignes géodésiques des surfaces à courbures opposées; il a donné une solution complète de ce problème dans un memoire du plus haut intérêt. Mais ce n'est pas aux géodésiques des surfaces à courbures opposées que les trajectoires du problème des trois corps sont comparables, c'est au contraire aux géodésiques des surfaces convexes.

J'ai donc abordé l'étude des lignes géodésiques des surfaces convexes; malheureusement le problème est beaucoup plus difficile que celui qui a été résolu par M. HADAMARD. J'ai donc dû me borner à quelques resultats partiels, relatifs surtout aux géodésiques fermées qui jouent ici le rôle des solutions périodiques du problème des trois corps.

* Presented to the Society at the St. Louis meeting, September 17, 1904. Received for publication January 4, 1905.

§ 2. *Foyers et caustiques.*

Soit S une surface *convexe, analytique;* supposons que ses deux rayons de courbure principaux restent constamment compris entre deux limites L_1 et L_2 et par conséquent sa courbure totale entre L_1^2 et L_2^2. Nous voyons d'abord que le rayon de courbure de l'une quelconque de ses géodésiques restera toujours compris entre L_1 et L_2. Je commence par rappeler les propriétés essentielles des *foyers* et des *caustiques.*

Soit O un point fixe de la surface S. Envisageons une géodésique OM passant par le point O, soit OH une autre géodésique fixe passant par ce même point O, soit v l'angle sous lequel ces deux géodésiques se coupent en O, soit u l'arc OM compté sur le géodésique. Ces deux quantités u et v peuvent être regardées comme des coordonnées définissant la position du point M sur la surface; ce sont, en quelque sorte, des coordonnées polaires, le point O jouant le rôle du pôle, et la géodésique OH celui de l'axe polaire. Le carré de l'élément d'arc sera de la forme

$$ds^2 = du^2 + \lambda^2 dv^2,$$

où λ est une fonction de u et de v.

Il est à remarquer qu'un point quelconque M de la surface correspond à une infinité de couples de valeurs des coordonnées u et v, et en effet il y a une infinité de manières d'aller de O en M en suivant une géodésique. Par exemple si S se réduit à une sphère de rayon 1, on a

$$ds^2 = du^2 + \sin^2 u \, . \, dv^2$$

et les couples de valeurs

$$(u, v), \ (u + 2\pi, v), \ (u + 4\pi, v), \cdots,$$

correspondent à un même point M.

Considérons deux géodésiques issues du point O et infiniment peu différentes, et supposons qu'elles se coupent de nouveau en un point M. Nous aurons pour le point d'intersection deux couples de valeurs des coordonnées et l'on obtiendra l'un ou l'autre suivant qu'on atteindra ce point M en suivant la première géodésique, ou bien en suivant la seconde qui en diffère fort peu. Comme ces deux géodésiques sont infiniment peu différentes, les deux couples de valeurs seront très peu différents, soit (u, v), et $(u + du, v + dv)$. Comme le point M représenté par ces deux couples est le même, on aura

$$ds^2 = du^2 + \lambda^2 dv^2 = 0,$$

soit

$$du = 0, \ \lambda = 0.$$

Si donc nous envisageons deux géodésiques infiniment voisines issues du point O, elles se recouperont successivement en une infinité de points d'après un

théorème de M. HADAMARD et les points d'intersection successifs pourront être désignés par $F_1, F_2, \cdots, F_n$. Le point F_n est ordinairement appelé le $n^{\text{ème}}$ *foyer* du point O; on voit que F_p est le $(p-q)^{\text{ème}}$ foyer de F_q.

On appelle *caustique* l'enveloppe de toutes les géodésiques issues du point O; et on voit que la caustique est le lieu des foyers du point O. Cette caustique d'après ce qui précède a pour équation

$$\lambda = 0.$$

Considérons un instant v comme une constante, de façon à suivre constamment une des géodésiques issues de O, et désignons par λ', λ'', etc., les dérivées successives de λ par rapport à u. On sait que

$$-\frac{\lambda''}{\lambda}$$

représente la courbure totale de sorte que l'on a

$$\frac{1}{L_1^2} < -\frac{\lambda''}{\lambda} < \frac{1}{L_2^2}. \tag{1}$$

On conclut aisément de ces inéqualités:

1°. que λ, λ' et λ'' sont toujours finis.

2°. que les racines de l'équation $\lambda = 0$ sont séparées par celles de l'équation $\lambda' = 0$ et réciproquement.

3°. que si u_0 est une racine de l'equation $\lambda = 0$ et u_1 une racine de l'equation $\lambda' = 0$; si entre u_0 et u_1 il n'y a aucune racine ni de $\lambda = 0$, ni de $\lambda' = 0$, on aura les inégalités

$$\frac{\pi}{2} L_2 < |u_1 - u_0| < \frac{\pi}{2} L_1, \tag{2}$$

Si donc nous considérons u et v comme les coordonnées *polaires* d'un point dans un plan, et que nous construisions dans ce plan les courbes $\lambda = 0$, $\lambda' = 0$, nous voyons que ces courbes se composent d'une série d'ovales fermées s'enveloppant mutuellement et enveloppant le pôle, de sorte que si l'on s'éloigne du pôle en suivant un rayon vecteur on rencontre alternativement une ovale appartenant à $\lambda = 0$ et une ovale de $\lambda' = 0$; la distance de deux ovales consécutives est comprise entre $\pi/2L_2$ et $\pi/2L_1$. On a d'ailleurs pour $u = 0$,

$$\lambda = 0, \ \lambda' = 1.$$

de sorte que l'on peut considérer le pôle comme l'une des ovales de la courbe $\lambda = 0$, cette ovale se réduisant à un seul point.

Les 1^{ers} foyers se trouvent sur la $1^{\text{ère}}$ ovale, les 2^{das} foyers sur la 2^{de}, et ainsi de suite. On voit donc que les foyers des différents ordres sont entièrement séparés les uns des autres. C'est ce qui nous permet de définir la $n^{\text{ème}}$ *caustique* comme le lieu des $n^{\text{èmes}}$ *foyers*.

Nous appellerons P le plan sur lequel nous représentons ainsi les points de S, en prenant pour coordonnées polaires u et v. Dans ce plan, nous appellerons $C_1, C_2, \cdots$ les ovales successives de la courbe $\lambda = 0$, (C_0 se réduisant au pôle); nous appellerons $C'_1, C'_2, \cdots$, les ovales successives de la courbe $\lambda' = 0$. Dans le cas ou la surface S serait de révolution, le point O étant le pôle, λ ne dépendrait que de u et les ovales C et C' se réduiraient à des cercles concentriques. Soient x, y, z les coordonnées du point M et

$$F(x, y, z) = 0$$

l'équation de la surface S. Les équations d'une ligne géodésique peuvent s'écrire

$$\frac{d^2x}{du^2} = \mu \frac{dF}{dx}, \qquad \frac{d^2y}{du^2} = \mu \frac{dF}{dy}, \qquad \frac{d^2z}{du^2} = \mu \frac{dF}{dz}, \tag{3}$$

μ étant déterminé par l'équation

$$\Sigma \frac{d^2F}{dx^2}\left(\frac{dx}{du}\right)^2 + 2\Sigma \frac{d^2F}{dxdy}\frac{dx}{du}\frac{dy}{du} + \mu\Sigma\left(\frac{dF}{dx}\right)^2 = 0. \tag{4}$$

Les valeurs initiales de x, y, z seront les coordonnées du point O; quant aux valeurs initiales de $dx/du, dy/du, dz/du$, elles dépendront linéairement de $\cos v$ et $\sin v$.

Les équations différentielles (3), (4) ne peuvent admettre de point singulier réel; elles n'en admettraient en effet que si μ devenait infini, c'est à dire si $\Sigma(dF/dx)^2$ s'annulait, c'est à dire si l'on avait à la fois

$$\frac{dF}{dx} = \frac{dF}{dy} = \frac{dF}{dz} = 0,$$

c'est à dire si la surface S avait un point singulier, ce qui n'est pas. Nous sommes donc certain que la solution envisagée de nos équations différentielles n'ira passer par aucun point singulier. Je puis donc lui appliquer le théorème que j'ai démontré aux Nos. 23 à 28 du Tome 1er de *Méthodes nouvelles de la Mécanique céleste*. Il en résulte que x, y et z sont des fonctions holomorphes de u et des valeurs initiales, et par conséquent de u et de v; je veux dire que x, y, z sont développables suivant les puissances de $u - u_0$, $v - v_0$, pourvu que $|u - u_0|$ et $|v - v_0|$ soient assez petits et cela quelles que soient les valeurs de u_0 et v_0.

Revenons à nos caustiques; soient Q_1 et Q_2 deux points d'une caustique entre lesquels nous supposerons qu'il n'y ait aucun point singulier. Quelle sera (sur le surface S) l'arc de caustique compris entre ces deux points? Si u_1 et u_2 sont les valeurs de u correspondant à ces deux points; cet arc est

$$|u_1 - u_2|.$$

On le verrait par un raisonnement tout pareil à celui qui donne l'arc de la développée d'une courbe plane.

L'élément d'arc de caustique est donc égal à du.

Reprenons les deux géodésiques infiniment voisines issues du point O et correspondant aux angles v et $v + dv$. La distance d'un point de la 1ère géodésique à la 2de sera évidemment λdv; quand on approchera d'un foyer, cette distance tendra vers 0 puisque λ tendra vers 0, les deux géodésiques se recouperont sous un angle $\lambda' dv$. L'angle de contingence géodésique de la caustique est donc $\lambda' dv$. La courbure géodésique de la caustique est donc

$$\frac{\left(\frac{d\lambda}{du}\right)^2}{\left(\frac{d\gamma}{dv}\right)}.$$

Cela posé reprenons l'équation $\lambda = 0$; en aucun point on ne pourra avoir $\lambda' = 0$: les deux courbes $\lambda = 0$, $\lambda' = 0$ ne peuvent se couper, puisque la distance des deux ovales C et C' est au moins égale à $\pi/2L_2$; donc l'équation $\lambda = 0$ peut être résolue par rapport à u, et u est une fonction holomorphe de v. Comme x, y, z sont fonctions holomorphes de u et de v, nous voyons que *dans l'équation de la caustique les coordonnées x, y, z sont des fonctions holomorphes de v.*

Toutes les fois que $d\lambda/dv$ n'est pas nul, l'équation $\lambda = 0$ peut être résolue par rapport à v, et v est fonction holomorphe de u. *Les coordonnées d'un point de la caustique sont donc des fonctions holomorphes de l'arc u.* Tous les points singuliers de notre caustique nous seront donc donnés par l'équation $d\lambda/dv = 0$. J'ajoute que la courbure géodésique de la caustique ne peut pas s'annuler ni changer de signe, puisque $d\lambda/du$ ne peut s'annuler. Il n'y a donc rien qui corresponde à une sorte de point d'inflexion.

Les points singuliers de la caustique correspondent donc aux racines de l'équation $d\lambda/dv = 0$, c'est à dire aux minima de u quand on décrit l'une des ovales C dans le plan P.

Pour nous rendre compte de la nature de ces points singuliers prenons un instant pour origine le point singulier, pour plan des xy le plan tangent, pour axe des x la tangente à la courbe, et choisissons l'origine de l'angle v de façon que $v = 0$ au point singulier. Alors x et y sont développables suivant les puissances de v, le 1er terme du développement est en v^m pour x, en v^n pour y; et les entiers m et n satisfont à l'inégalite

$$m < n.$$

Si λ s'annule ainsi que ses p premières dérivées par rapport à v et que la

$(p+1)^{\text{ème}}$ ne s'annule pas, comme d'ailleurs $d\lambda/du$ ne s'annule pas; on voit que le développement de u commence par un terme en v^{p+1}. On aura donc

$$m = p + 1.$$

D'autre part $d\lambda/dv$ commence par un terme en v^p de sorte que pour un point très voisin du point singulier, la courbure géodésique est de l'ordre de v^{-p}.

Or cette courbure a pour expression

$$\frac{\frac{dx}{dv}\frac{d^2y}{dv^2} - \frac{dy}{dv}\frac{d^2x}{dv^2}}{\left[\left(\frac{dx}{dv}\right)^2 + \left(\frac{dy}{dv}\right)^2\right]^{\frac{3}{2}}};$$

elle est donc de l'ordre de

$$m + n - 3 - 3(m-1) = n - 2m.$$

On a donc

$$n - 2m = -p,$$

d'où finalement

$$m = p + 1, \qquad n = p + 2.$$

Le cas le plus simple est celui de

$$p = 1, \qquad m = 2, \qquad n = 3;$$

le point singulier est alors un point de rebroussement ordinaire.

Ces points de rebroussement correspondent aux maxima et aux minima de u, mais les deux cas doivent être distingués. S'il s'agit d'un minimum, et si R est le point de rebroussement en question, les deux branches de la caustique touchent non la géodésique OR, mais son prolongement; s'il s'agit d'un maximum, les deux branches de la caustique sont dirigées de R vers O.

Nous sommes ainsi amenés à distinguer 4 sortes de foyers.

1°. Les *foyers ordinaires*, correspondant aux points non singuliers de la caustique ($p = 0$).

2°. Les *foyers en pointe*, correspondant aux points de rebroussement ordinaires ($p = 1$), qui sont des minima de u.

3°. Les *foyers en talon*, correspondant aux points de rebroussement ordinaires, ($p = 1$) qui sont des maxima de u.

Ces deux expressions sont empruntées à l'art des chemins de fer; je me représente une trifurcation, où nous avons la voie principale représenté par la géodésique OR et deux embranchements représentés par les deux branches de la caustique; un train qui suivrait la ligne principale en allant de O vers R rencontrerait les aiguilles en pointe dans le 1$^{\text{er}}$ cas et en talon dans le 2$^{\text{d}}$.

4°. Les *foyers singuliers*, correspondant aux points singuliers d'ordre plus élevé ($p > 1$). On voit pour ce qui précède que ces points singuliers rentrent dans tous les cas dans une classe très particulier.

Nous introduisons maintenant la notion de *ligne de partage*. Supposons que l'on joigne chaque point de la surface S au point O par le plus court chemin. Ce plus court chemin sera un arc de géodésique. Si $OAPM$ est le plus court chemin de O à M; si le point P est situé sur l'arc $OAPM$, l'arc OAP sera manifestement le plus court chemin de O à P. Il suit de là que le plus court chemin de O à M ne peut jamais croiser le plus court chemin de O à M'.

Soit OM une géodésique quelconque passant par O; on pourra trouver sur cette géodésique un point P, tel que le plus court chemin de O à un point Q situé sur la géodésique OM entre O et P, soit précisément l'arc OQ de cette géodésique OM, mais que cela ne soit plus vrai si le point Q est au delà de P. On dit alors que P est *l'extrémité* du plus court chemin OP.

Nous pouvons alors conclure que par tout point de S passe un plus court chemin et un seul. Il y a exception pour les extrémités des divers plus courts chemins; si P est l'une de ces extrémités; du point P partiront *au moins deux* plus courts chemins qui auront l'un et l'autre leur extrémités en P. Tous les plus courts chemins qui ont leur extrémité en P ont même longueur.

Le lieu des points qui sont les extrémités de deux on plusieurs plus courts chemins forment un ensemble de lignes que l'on peut appeler *lignes de partage*, Si un point P partent seulement deux plus courts chemins, par le point P passe une seule ligne de partage dont la tangent est la bissectrice de l'angle formé par les tangentes aux deux plus courts chemins.

Si du point P partent plus de deux plus courts chemins, au point P aboutissent plusieurs lignes de partage, dont les tangentes sont les bissectrices des angles formés par les tangentes à deux plus courts chemins consécutifs.

L'ensemble des lignes de partage ne divise pas la surface S en deux régions, puisque l'on peut aller du point O à un point quelconque M de la surface sans traverser aucune ligne de partage; il suffit pour cela d'aller de O en M par le plus court chemin.

L'ensemble des lignes de partage, ou une partie de ces lignes, ne peut donc jamais constituer un polygone fermé; il formera une sorte de *système rameux*. où les bifurcations seront représentées par les points où aboutissent plus de deux plus courts chemins. Que représenteront alors les extrémités des rameaux? Supposons que nous suivions une ligne de partage PQR et que le point R soit l'extrémité de cette ligne. Du point Q partiront deux plus courts chemins de même longueur; quand le point Q ira de P en R, ces deux plus courts chemins varieront d'une manière continue. Au point R ils devront se confondre en un seul.

Il est aisé de voir que les points R sont les seuls points des lignes de partage qui se trouvent sur la 1ère caustique; (je veux dire bien entendu qu'en aucun autre point d'une ligne de partage, l'un des plus courts chemin qui y'aboutissent

ne touche la 1ère caustique). On voit d'ailleurs que ces points R ne peuvent être des foyers ordinaires, mais seulement des foyers en pointe.

§ 3. *Géodésiques d'un sphéroïde.*

Cherchons les géodésiques d'une surface très peu différente d'une sphère; nous n'aurons pour cela qu'à appliquer la méthode de la variation des constantes de LAGRANGE. Sur la sphère, les lignes géodésiques sont les grands cercles et l'on peut adopter comme éléments définissant le mouvement d'un point sur ce grand cercle de la même manière que les éléments elliptiques définissent le mouvement Képlerien d'une planète, les quatre quantités suivantes: les coordonnées du pôle du grand cercle, la vitesse uniforme du mouvement ou une fonction de cette vitesse, la longitude du point mobile sur ce grand cercle, comptée à partir d'une certaine origine. Mais il convient d'abord de mettre les équations sous la forme canonique. Si nous mettons l'élément linéaire d'une surface quelconque sous la forme

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2,$$

on aura pour l'expression T de la demi-force vive qui figure dans les équations de Lagrange ou de Hamilton

$$T = \tfrac{1}{2}(Eu'^2 + 2Fu'v' + Gv'^2),$$

d'où

$$U = \frac{dT}{du'} = Eu' + Fv',$$

$$V = \frac{dT}{dv'} = Fu' + Gv',$$

et

$$T = \tfrac{1}{2}(\mathfrak{E}U^2 + 2\mathfrak{F}UV + \mathfrak{G}V^2),$$

où $\mathfrak{E}$, $\mathfrak{F}$ et $\mathfrak{G}$ ont des valeurs faciles à calculer. Dans la cas de la sphère on a simplement

$$T = \tfrac{1}{2}(u'^2 + \sin^2 u \cdot v'^2),$$

d'où

$$T = \tfrac{1}{2}\left(U^2 + \frac{1}{\sin^2 u}V^2\right).$$

Dans le cas d'une surface très peu différente de la sphère, on aura

$$T = T_0 + \mu T_1,$$

où

$$T_0 = \tfrac{1}{2}\left(U^2 + \frac{1}{\sin^2 u}V^2\right),$$

où μ sera très petit et où

$$T_1 = \tfrac{1}{2}(eU^2 + 2fUV + gV^2)$$

e, f et g étant des fonctions quelconques de u et de v.

Formons l'équation de Jacobi.

Supposons d'abord $\mu = 0$ de façon que notre sphéroïde se réduise à une sphère.

Soit θ la longitude du noeud de l'orbite circulaire décrite par le point sur la sphère, sur le plan de l'équateur; soit i l'inclinaison de cette orbite; soit λ la longitude du point mobile sur la sphère comptée à partir du noeud; soit ω la vitesse de circulation; (le rayon de la sphère est pris pour unité ainsi que la masse du point mobile); alors la demi-force vive sera $\omega^2/2$; la constante des aires sera ω dans le plan de l'orbite et $\omega \cos i$ dans le plan de l'equateur.

Il est aisé d'exprimer u, v, U, V en fonctions de θ, λ, i, ω, et on verrait comme en mécanique céleste que, en posant $G = \omega \cos i$,

$$\omega d\lambda + G d\theta - U du - V dv$$

est une differentielle exacte.

Si nous ne supposons plus $\mu = 0$, nous pourrons toujours considérer des variables nouvelles θ, λ, i, ω, liées à u, v, U et V par les mêmes relations que dans le cas de la sphère et que nous considérerons comme "les éléments de l'orbite osculatrice." Les équations conservent alors leur forme canonique et peuvent s'écrire

$$\frac{d\lambda}{dt} = \frac{dT}{d\omega}, \qquad \frac{d\omega}{dt} = -\frac{dT}{d\lambda}, \qquad \frac{dG}{dt} = -\frac{dT}{d\theta}, \qquad \frac{d\theta}{dt} = \frac{dT}{dG}.$$

Quelle est la forme de T? Observons d' abord que $T_0 = \omega^2/2$. Remarquons en outre que u, v, u'/ω, v'/ω, U/ω, V/ω et par conséquent T/ω^2 ne dépendent que de i, λ, θ. Posons donc

$$\omega t = \tau, \quad T = \omega^2 S, \quad T_0 = \omega^2 S_0, \quad T_1 = \omega^2 S_1;$$

nos équations deviendront

$$\frac{d\omega}{\omega d\tau} = -\frac{dS}{d\lambda}, \qquad \frac{di}{d\tau}\sin i = \frac{dS}{d\theta} - \frac{dS}{d\lambda}\cos i,$$

$$\frac{d\lambda}{d\tau} = 2S + \frac{dS}{di}\operatorname{cotg} i, \qquad \frac{d\theta}{d\tau} = -\frac{1}{\sin i}\frac{dS}{di},$$

qui admettent comme il convient l'intégrale $\omega^2 S = \text{const.}$

Comme S_0 se réduit à $1/2$ nous pouvons écrire

$$\frac{di}{d\tau} = \frac{\mu}{\sin i}\frac{dS_1}{d\theta} - \mu \operatorname{cotg} i \frac{dS_1}{d\lambda}, \qquad \frac{d\theta}{d\tau} = -\frac{\mu}{\sin i}\frac{dS_1}{di},$$

$$\frac{d\lambda}{d\tau} = 1 + 2\mu S_1 + \mu \operatorname{cotg} i \frac{dS_1}{di}.$$

Comme μ est très petit, nous pouvons dans les seconds membres remplacer les inconnus par leurs valeurs approchées qui sont celles qui correspondent au cas

de la sphère; c'est à dire des valeurs constantes pour i et θ, pour λ, la valeur approchée $\tau + \text{const.}$

Au point de vue qui nous occupe, il sera d'ailleurs préférable de prendre λ pour variable indépendante et d'écrire

$$\frac{di}{d\lambda} = \frac{\dfrac{\mu}{\sin i}\dfrac{dS_1}{d\theta} - \mu \operatorname{cotg} i \dfrac{dS_1}{d\lambda}}{1 + 2\mu S_1 + \mu \operatorname{cotg} i \dfrac{dS_1}{di}},$$

ou en négligeant μ^2,

$$\frac{di}{d\lambda} = \frac{\mu}{\sin i}\frac{dS_1}{d\theta} - \mu \operatorname{cotg} i \frac{dS_1}{d\lambda}, \tag{1}$$

et de même

$$\frac{d\theta}{d\lambda} = -\frac{\mu}{\sin i}\frac{dS_1}{di}. \tag{2}$$

Dans les 2^{ds} membres, on conservera alors λ et on remplacera i et θ par leurs valeurs approchées qui sont des constantes.

Rendons-nous compte de la signification géométrique de cette fonction S_1; comment d'abord ferons-nous correspondre les points du sphéroïde à ceux de la sphère? Le choix de cette correspondance est arbitraire dans une assez large mesure. Le plus simple est de faire correspondre les points où le plan tangent a même direction, c'est adopter ce qu'on appelle la représentation sphérique des surfaces. Dans ces conditions, si l'on veut me permettre le langage de la Géodésie, u et v représenteront, sur le sphéroïde la colatitude et la longitude *astronomiques*.

Considérons alors sur le sphéroïde deux points infiniment voisins M et M', et sur la sphère les deux points correspondants M_1 et M'_1. Soient u et v les coordonnées de M (ou de M_1); soient $u + u'dt$, $v + v'dt$ celles de M' (ou de M'_1). Soit ds l'arc MM' et ds_1 l'arc $M_1M'_1$, nous voyons que le rapport ds/ds_1 depend non-seulement de la position du point M, mais de l'orientation de l'arc MM'; lorsque cet arc appartient à une ligne de courbure, il n'est pas autre chose que le rayon de courbure principal correspondant.

Si M et M' représentent deux positions d'un point mobile sur le sphéroïde aux instants t et $t + dt$, on aura pour l'énergie cinétique

$$T = \tfrac{1}{2}\left(\frac{ds}{dt}\right)^2.$$

Il est manifeste que les deux couples de points MM' et $M_1M'_1$ correspondent aux mêmes valeurs de u, v, u', v'; mais ils ne correspondent pas aux mêmes valeurs de u, v, U, V, ni par conséquent de i, θ, ω, λ; car la relation qui lie U et V à u, v, u', v', n'est pas la même pour $\mu = 0$ c'est à dire pour la sphère, et pour $\mu > 0$, c'est à dire pour le sphéroïde. Soit alors sur la sphère un couple

de points infiniment voisins $M_1 M_2'$, choisis de façon à correspondre aux mêmes valeurs de u, v, U, V (ou de i, θ, ω, λ) que le couple de points MM' du sphéroïde.

Soient alors u et v les coordonnées du point M_1, $u + u_0' dt$, $v + v_0' dt$ celles du point M_2'; soit ds_2 l'arc $M_1 M_2'$. Alors i et θ désigneront la colatitude et la longitude du pôle du grand cercle qui passe par M_1 et M_2'; λ sera l'arc de ce grand cercle compris entre le point M_1 et l'équateur, et on aura enfin

$$M_1 M_2' = ds_2 = \omega dt.$$

Soit

$$T_0 = \tfrac{1}{2}(u'^2 + v'^2 \sin^2 u),$$

$$T_0' = \tfrac{1}{2}(u_0'^2 + v_0'^2 \sin^2 u),$$

on aura évidemment,

$$T_0 = \tfrac{1}{2}\left(\frac{ds_1}{dt}\right)^2, \qquad T_0' = \tfrac{1}{2}\left(\frac{ds_2}{dt}\right)^2 = \frac{\omega^2}{2}.$$

Si nous désignons de même par T_1' ce que devient T_1 quand on y remplace u' et v' par u_0' et v_0', la différence $T_1' - T_1$ sera de l'ordre de μ de même que $dT_1'/du_0' - dT_1/du'$. Comme U et V doivent avoir même valeur pour le couple MM' sur la sphère et pour le couple $M_1 M_1'$ sur le sphéroïde, on aura

$$U = \frac{dT_0'}{du_0'} = \frac{dT}{du'},$$

$$V = \frac{dT_0'}{dv_0'} = \frac{dT}{dv'},$$

d'où en négligeant μ^2

$$\frac{dT_0'}{du_0'} = \frac{dT_0}{du'} + \mu \frac{dT_1'}{du_0'},$$

$$\frac{dT_0'}{dv_0'} = \frac{dT_0}{dv'} + \mu \frac{dT_1'}{dv_0'}.$$

Je multiplie la 1ère par u_0'; la 2de par v_0' et j'ajoute; j'obtiens alors par les propriétés des formes quadratiques,

$$\Sigma u_0' \frac{dT_0'}{du_0'} = \Sigma u_0' \frac{dT_0}{du'} + \mu \Sigma u_0' \frac{dT_1'}{du_0'} = u' \frac{dT_0'}{du_0'} + 2\mu T_1',$$

d'où

$$\Sigma (u_0' - u') \frac{dT_0'}{du_0'} = 2\mu T_1'.$$

D'autre part, nous pouvons développer T_0 par la formule de Taylor suivant les puissances croissantes de $u' - u_0'$, $v' - v_0'$; nous avons alors en négligeant μ^2, c'est à dire en négligeant les termes du 2d ordre,

$$T_0 = T_0' + \Sigma \frac{dT_0'}{du_0'}(u' - u_0') = \frac{\omega^2}{2} \cdot 2\mu T_1',$$

ou toujours en négligeant μ^2,

$$T_0 = \frac{\omega^2}{2} - 2\mu T_1,$$

d'où

$$T = \frac{\omega^2}{2} - \mu T_1.$$

Des deux relations

$$T = T'_0 - \mu T_1, \qquad T = T_0 + \mu T_1$$

on tire d'abord

$$2T = T_0 + T'_0,$$

c'est à dire

$$2ds^2 = ds_1^2 + ds_2^2,$$

ou en négligeant μ^2,

$$2ds = ds_1 + ds_2.$$

Ce que nous appelons S_1, c'est le rapport $(T - \omega^2)/\mu\omega^2$, c'est à dire

$$(ds^2 - ds_2^2)/\mu ds_2^2,$$

ou en négligeant les puissances supérieures de μ

$$S_1 = \frac{ds_1^2 - ds^2}{\mu ds^2} = \frac{2}{\mu}\frac{ds_1 - ds}{ds} = \frac{2}{\mu}\frac{ds_1 - ds}{ds_1}.$$

Telle est la signification géométrique de S_1.

Reprenons maintenant les équations (1) et (2). Les 2^{ds} membres sont des fonctions périodiques de λ. Pour que i et θ soient également des fonctions périodiques de λ, c'est à dire pour que la géodésique soit fermée il faut et il suffit que la *valeur moyenne* de ce 2^d membre, considéré comme fonction périodique de λ, soit nulle.

Quelle est cette valeur moyenne? Développons S_1 suivant les cosinus et les sinus des multiples de λ et soit R la valeur moyenne de S_1, c'est à dire l'ensemble des termes indépendants de λ. Il est clair que dR/di sera la valeur moyenne de ds_1/di, et $dR/d\theta$ celle de $dS_1/d\theta$. Quant à la valeur moyenne de $dS_1/d\lambda$, elle sera nulle.

Pour que les valeurs moyennes de nos 2^{ds} membres soient nulles, il faut donc et il suffit que

$$\frac{dQ}{di} = \frac{dR}{d\theta} = 0.$$

Les géodésiques fermées répondront donc aux maxima, aux minima, et aux minimax de la fonction R. A l'exemple des Anglais, j'appelle minimax les points où s'annulent les dérivées du 1er ordre d'une fonction de deux variables sans qu'on ait ni maximum, ni minimum.

Il faut d'abord rechercher la signification géométrique de la fonction R. Si nous donnons à i et à θ des valeurs constantes et que nous fassions varier λ,

comment variera le point M. Le point M_1 décrira sur la sphère un grand cercle, le plan tangent au point M restera donc parallèle à une droite D dont la direction est définie par la colatitude i et la longitude θ. Construisons un cylindre circonscrit au sphéroïde et dont les génératrices sont parallèles à D; il touche le sphéroïde suivant une courbe fermée C. *Le point M restera sur cette courbe C.* Cette courbe C sera ce qu'on pourrait appeler, dans le langage de la géodésie, un grand cercle astronomique.

On a alors

$$R = \frac{1}{2\pi}\int_0^{2\pi} S_1 \, d\lambda.$$

Soient M, M' deux points infiniment voisins de la courbe C, soit M_1, M_1' les deux points correspondants de la sphère. Soit comme plus haut

$$MM' = ds, \qquad M_1 M_1' = ds_1.$$

On aura $ds_1 = d\lambda$ et

$$S_1 = \frac{2}{\mu}\frac{ds_1 - ds}{ds_1},$$

d'où

$$R = \frac{1}{2\pi}\int \frac{2}{\mu}\frac{ds_1 - ds}{d\lambda}\, d\lambda = \frac{1}{\pi\mu}(s_1 - s).$$

Ici $s_1 = 2\pi$ est la longueur totale du grand cercle de la sphère, et s est la longueur totale de la courbe C.

Ainsi les maxima, minima et minimax de R correspondent aux minima, maxima et minimax de la longueur totale des courbes C.

Considérons sur la sphère le grand cercle qui correspond à une courbe C, soient P et P' les deux pôles de ce grand cercle; la colatitude et la longitude du point P sont i et θ; et R est une fonction de i et de θ. Remarquons d'abord que cette fonction a même valeur aux deux points diamétralement opposés P et P'. Construisons les courbes $R = \text{const.}$, c'est à dire les courbes qui joignent les différents points P où la fonction R a la même valeur. Nous aurons une série de courbes telles que par chaque point de la sphère passe une de ces courbes et une seule. Aux minima et aux maxima de R correspondent des points isolés de ces courbes, aux minimax des points doubles à tangentes réelles. Un théorème d'Analysis Situs nous apprend que le nombre total des minima et des maxima, c'est à dire des points isolés, surpasse de 2 unités celui des minimax, c'est à dire des points doubles. Je me bornerai à renvoyer à ce sujet à mon mémoire sur les courbes définies par les équations différentielles (Journal de Liouville, sér. 3, tome 7).

Le nombre total des minima, maxima et minimax est donc un multiple de 4, plus 2. Mais à chaque courbe C correspondent deux pôles P et P' diamétralement opposés, donc à chaque géodésique fermée correspondent *deux* minima,

maxima ou minimax. *Le nombre total des géodésiques fermées est donc impair*, puisque c'est la moitié du nombre total des minima, maxima et minimax.

Il ne s'agit bien entendu ici que des géodésiques fermées qui subsistent *quelque petit que soit* μ. Il peut y en avoir et il y en a une infinité d'autres. Une question reste à traiter. L'existence des géodésiques fermées dont nous venons de parler à été établie par un calcul approximatif puisque nous avons négligé μ^2. Cette existence peut-elle être démontrée rigoureusement? Ou plutôt cette démonstration rigoureuse ne peut elle pas se tirer directement de notre calcul approximatif? Oui, cela peut se faire et par un procédé connu; on n'a qu'à appliquer les principes exposés dans le Chapitre 3 du Tome 1 de mon ouvrage, *Les Méthodes nouvelles de la Mécanique céleste.*

§ 4. *Le principe de continuité analytique.*

Soient S et S' deux surfaces quelconques satisfaisant aux conditions que nous nous sommes imposées dans les paragraphes précédents et que j' appellerai pour abréger les conditions A. *On peut toujours passer de l'une à l'autre d'une manière continue.* Je veux dire que l'on peut supposer une surface variable Σ qui ayant pour équation générale

$$F(x, y, z, t) = 0,$$

où le 1er membre F est une fonction *analytique* des coordonnées x, y, z et d'un paramètre variable t (qui restera analytique au moins pour tous les systèmes de valeurs de x, y, z, t, tels que 1°. t soit compris entre 0 et 1; 2°. x, y, z soient réels 3°. F soit nul.) et de telle façon: 1°. Que cette surface Σ se réduit à S pour $t = 0$ et à S' pour $t = 1$; 2°. Que quand t varie d'une manière continue depuis 0 jusqu'à 1, cette surface Σ ne cesse jamais de satisfaire aux conditions A.

Ainsi la surface Σ quand t variera d'une manière continue entre 0 et 1, restera convexe et analytique, ses rayons de courbure restent compris entre des limites déterminées. Considérons sur cette surface variable Σ une ligne géodésique satisfaisant à une certaine condition, par exemple à celle d'être fermée. Soit $y = \phi(x)$, $z = \psi(x)$ l'équation de cette géodésique. Définissons cette géodésique, par les données initiales; appelons par exemple y_0 et z_0 les valeurs de y et de z pour $x = 0$, de sorte que

$$y_0 = \phi(0), \qquad z_0 = \psi(0).$$

Soient de même y'_0, z'_0 les valeurs initiales de dy/dx et dz/dx, de telle façon que

$$y'_0 = \phi'(0), \qquad z'_0 = \psi'(0).$$

Ces données initiales y_0, z_0, y'_0, z'_0 suffisent pour distinguer notre géodésique de toutes les autres géodésiques tracées sur la surface Σ.

Cela posé, si nous écrivons que cette géodésique est fermée, nous obtiendrons

entre y_0, z_0, y'_0, z'_0 certaines relations qui seront analytiques. Donc y_0, z_0, y'_0, z'_0 sont liés à t par des relations analytiques; si nous considérons y_0, z_0, y'_0, z'_0, t comme les coordonnées d'un point dans l'espace à 5 dimensions, ces relations définiront une certaine courbe analytique que j'appelle C. Nous envisagerons une branche de cette courbe analytique que j'appelle B. Il est aisé de définir cette branche et de la distinguer des autres par continuité. Dans le voisinage de $t = t_0$, les quantités y_0, y'_0, z_0, z'_0 sont susceptibles d'un ou de plusieurs développements et chacun de ces développements procède suivant les puissances de $(t - t_0)^{1/q}$; la branche B se composera de tous les points réels de la courbe C qui correspondent à l'un de ces développements; quand le développement ne sera plus valable, on la poursuivra par le procédé de la continuation analytique. Si q est impair, cette branche comprendra un point et un seul, tant pour $t = t_0 - e$ que pour $t = t_0 + e$; si q est pair, elle comprendra deux points pour $t = t_0 - e$ et zéro pour $t = t_0 + e$; ou inversement.

Nous dirons que les différentes géodésiques fermées qui appartiennent aux différents points d'une même branche B font partie d'une même *série continue*.

Considérons les différents géodésiques fermées d'une même série continue qui appartiennent à une même surface Σ. Elles correspondent aux points d'intersection de la branche B avec le plan $t = t_0$. Le nombre de ces points ne peut varier que quand t_0 prend une valeur telle que le nombre que nous appelions plus haut q soit pair. Dans ce cas, ce nombre varie de deux unités. D'où cette conséquence:

Le nombre des géodésiques d'une surface Σ *qui font partie de une, deux ou plusieurs séries continues déterminées est constamment pair ou constamment impair.*

Examinons maintenant les différentes sortes de courbes fermées que l'on peut tracer sur une surface convexe. Ces courbes au point de vue de l'Analysis Situs se répartiront en une infinité de *types* caractérisés par le nombre et la disposition des points doubles.

Considérons une série continue de géodésiques fermées; elles pourront appartenir à différents types; comment pourront elles passer d'une type à l'autre? D'abord *elles ne pourront jamais avoir de point de rebroussement;* car les géodésiques des surfaces satisfaisant aux conditions A n'ont aucun point singulier. Ensuite, il ne pourra jamais arriver que deux branches de cette géodésique viennent à se toucher. Car deux géodésiques ne peuvent se toucher sans se confondre, attendu qu'un point et la tangente en ce point sur une surface convexe détermine complètement une géodésique.

Il résulte de là que *dans une même série continue, le nombre des points doubles d'une géodésique fermée demeure constant.* Ce nombre en effet, pourrait varier de deux manières: 1°. Si un point double à tangentes réelles devenait un point isolé à tangentes imaginaires, mais alors il faudrait passer par un

point de rebroussement. 2°. Si deux points doubles réels devenaient imaginaires conjugués, mais alors il faudrait qu'au moment du passage, deux branches de courbe vinssent à se toucher.

Nous venons de voir que tout cela est impossible; un seul point reste à éclaircir. J'ai dit que deux géodésiques ne peuvent se toucher *à moins de se confondre*. Ne serait il pas possible que pour certaines valeurs de t, la géodésique fermée que nous envisageons *dégénère* parce que deux de ses parties se confondraient entre elles? Il est clair que cela est possible. Une courbe fermée d'un certain type, peut se réduire en dégénérant à une courbe fermée d'un autre type parcourue plusieurs fois. C'est ainsi qu'un limaçon de Pascal, peut, à la limite, se réduire à un cercle parcouru deux fois. Il n'y a pas de raison pour que cela n'arrive pas pour les géodésiques et en fait cela arrive.

Si une courbe fermée du type T, peut dégénérer en une courbe fermée du type T' parcourue plusieurs fois, nous dirons que le type T' est *subordonné* au type T.

Supposons donc que nous considérions une série continue de géodésiques fermées correspondant à la branche de courbe B définie plus haut: que M, M', M'' soient 3 points infiniment voisins de cette branche, que M' soit entre M et M'', qu'en M la géodésique appartienne an type T; qu'en M' elle se réduise à une courbe fermée du type T' parcourue plusieurs fois; qu'en M'' elle appartienne an type T'', qui pourra être d'ailleurs identique à T. Le type T' est ainsi subordonné tant à T qu'à T''. Quand on passera de M en M'' on peut se demander si le nombre des points doubles a varié.

Soit t_0 la valeur de t qui correspond an point M'; nous poserons $\tau = (t - t_0)^{1/q}$ en choisissant l'entier q de façon que y_0, z_0, y_0', z_0' soient développables suivant les puissances de τ. Considérons l'un des points doubles de notre géodésique, quand t tend vers t_0, ce point se rapprochera d'un certain point D de la géodésique singulière du type T'. Prenons ce point pour origine et choisissons les axes de façon que le plan des yz ne soit tangent à aucune des branches de géodésique qui passent par ce point. Dans ces conditions, nous pourrons mettre les équations d'une quelconque de nos branches de géodésique sous la forme suivante; y et z seront développables suivant les puissances de x et de τ. Considérons deux de ces branches et soient

$$y = \Sigma \tau^m \phi_m(x), \qquad z = \Sigma \tau^m \psi_m(x),$$
$$y = \Sigma \tau^m \phi_m'(x), \qquad z = \Sigma \tau^m \psi_m'(x),$$

leurs équations. Je suppose que ces deux branches se confondent pour $\tau = 0$, on a donc

$$\phi_0(x) = \phi_0'(x), \qquad \psi_0(x) = \psi_0'(x).$$

Cherchons les points doubles, c'est à dire les points d'intersection de nos deux branches. Ils nous seront donnés par l'équation

$$\Sigma\tau^m[\phi_m(x)-\phi'_m(x)]=0$$

(laquelle entraîne d'ailleurs l'autre équation

$$\Sigma\tau^m[\psi_m(x)-\psi'_m(x)]=0$$

en tenant compte de l'équation de la surface Σ).

Soit $\phi_p(x)-\phi'_p(x)$ le premier des coefficients $\phi_m(x)-\phi'_m(x)$ qui ne s'annule pas identiquement; quand τ tendra vers zero, les points doubles tendront vers des limites, qui seront les solutions de l'équation

$$\phi_p(x)-\phi'_p(x)=0.$$

Par l'hypothèse le point $x=0$ est une de ces limites; on a donc

$$\phi_p(0)-\phi'_p(0)=0.$$

Nos points doubles sont alors donnés par l'équation

$$\Sigma\tau^{m-p}[\phi_m(x)-\phi'_m(x)]=0$$

dont le 1er membre, développable suivant les puissances de x et de τ, s'annule pour $x=\tau=0$; on peut en tirer x développé suivant les puissances de τ ce qui prouve que x reste réel aussi bien pour $\tau>0$ que pour $\tau<0$, et par conséquent qu'un point double réel, par exemple au point M, ne peut devenir imaginaire au point M'.

Il n'y aurait d'exception que si la dérivée

$$\frac{d}{dx}[\phi_p(x)-\phi'_p(x)]$$

s'annulait pour $x=0$. Or cela n'arrive pas; cela voudrait dire que deux branches se coupent en deux points infiniment voisins, c'est à dire qu'un point se confond avec son propre foyer, ou peut en devenir aussi voisin que l'on veut en comptant les distances sur la géodésique. Or nous avons vu plus haut que la distance d'un point à son foyer est toujours au moins égale à $r/2L_2$.

Donc même dans ce cas, le nombre des points doubles ne peut varier.

La disposition des points doubles peut elle varier? Oui, mais uniquement de la manière suivante: supposons trois branches de courbe qui se coupent en 3 points doubles de façon à former un petit triangle; à la limite ce petit triangle se réduit à un point et les 3 points doubles se confondent en un seul point triple; ensuite les 3 points doubles se séparent de nouveau, et les trois branches forment de nouveau un triangle, mais dont la disposition est différente. Si la branche 1 parcourue dans un certain sens rencontrait dans la 1ère disposition, la branche 2 d'abord, puis le branche 3; dans la 2de disposition, elle rencontrera la branche 3 d'abord puis la branche 2.

Il est aisé de voir que si les branches de courbe au lieu de déterminer un triangle

déterminent un polygone de plus de 3 côtés; la disposition des points doubles ne pourra changer, car il faudrait que deux de ces branches devinssent tangentes. Lorsqu'on pourra passer de cette manière du type T au type T', nous dirons que ces deux *types sont équivalents.* Il y a des types qui ne sont équivalents à aucun autre, par exemple le type sans point double. Une courbe fermée à points doubles partagera la surface convexe en un certain nombre de polygones curvilignes dont quelques-uns pourront être des biangles ou des uniangles. Si aucun de ces polygones n'est un triangle, le procédé précédent ne pourra être appliqué et le type ne sera équivalent à aucun autre. Considérons maintenant deux séries continues de géodésiques; soient G la géodésique de la 1ère serie, G' celle de la 2de; le nombre des points d'intersection de G et de G' ne pourrait varier que si ces deux géodésiques devenaient tangentes l'une à l'autre, ce qui est impossible.

Le nombre des points d'intersection de deux géodésiques appartenant à deux séries continues déterminées est invariable.

Supposons que pour t voisin de t_0, on trouve que le nombre q, dont nous avons parlé plus haut et qui figure dans l'expression $(t-t_0)^{1/q}$ est pair pour la 1ère série continue, et égal à 1 pour la 2de série. Il arrivera alors que pour $t<t_0$, nous aurons 2 géodésiques G et G'' de la 1ère serie, et que pour $t>t_0$ nous n'en aurons plus. Au contraire nous aurons toujours une géodésique G' et une seule de la 2de série, quel que soit t. Nous pouvons poser

$$(t-t_0)^{1/q}=\tau$$

et faire varier τ d'une manière continue; alors dans la 1ère série nous passerons de la géodésique G à la géodésique G''; et dans la 2de série nous partirons de la géodésique G' et reviendrons finalement à cette même géodésique G'.

Donc le nombre des points d'intersection de G et de G' est le même que celui des points d'intersection de G et de G''.

Quant à la disposition relative des points doubles de G et de G' et de leurs points d'intersection, elle ne pourra varier que de la façon que j'ai dite plus haut à propos de la disposition des points doubles d'une seule géodésique.

Comme première application de ce principe; étudions les géodésiques fermées sans point double; ces géodésiques formeront évidemment une ou plusieurs séries continues. D'après ce qui précède ces séries continues ne pourront contenir que des géodésiques fermées sans point double.

Donc le nombre total des géodésiques fermées sans point double est toujours pair ou toujours impair. Or nous avons vu au § précédent que pour un sphéroïde ce nombre est impair. De plus on peut passer d'un sphéroïde à une surface convexe quelconque d'une manière continue.

Donc sur une surface convexe quelconque, il y a toujours au moins une géodésique fermée sans point double et il y en a toujours un nombre impair.

Par exemple pour un ellipsoïde nous avons les trois sections principales par les plans de symétrie.

§ 5. *Stabilité et instabilité.*

Etudions maintenant les géodésiques peu différentes d'une géodésique fermée (avec ou sans point double). Pour cela nous nous servirons du système de coordonnées suivant. Soit G_0 la géodésique fermée, O un point fixe de cette géodésique; prenons sur G_0 à partir du point O, une longueur $OP = u$; au point P menons une géodésique G_1 normale à G_0; sur G_1 à partir de P prenons une longueur $PM = v$; les coordonnées du point M seront u et v. L'élément d'arc sera donné par

$$ds^2 = dv^2 + h^2 du^2.$$

On voit que $h = 1$ et $dh/dv = 0$ pour $v = 0$, et quel que soit u. Les équations d'une géodésique seront

$$\frac{d}{du}\frac{v'}{\sqrt{v'^2 + h^2}} = h\frac{dh}{dv},$$

v' étant la dérivée de v par rapport à u; ou en supposant v très petit pour chercher les géodésiques très voisines de G_0,

$$v'' = h''v, \tag{1}$$

où h'' est la valeur de d^2h/dv^2 pour $v = 0$. Alors $-h''$, qui represente la courbure de la surface au point $(u, 0)$, est une fonction de u qui est d'ailleurs périodique puisque G_0 est fermée.

L'équation (1) est une équation linéaire à coefficients périodiques; elle admet deux solutions remarquables de la forme suivante:

$$v = e^{\alpha u}\phi(u), \qquad v = e^{-\alpha u}\psi(u), \tag{2}$$

$\phi(u)$ et $\psi(u)$ étant périodiques en u.

Quatre cas peuvent se présenter:

1°. Ou bien α est purement imaginaire et les deux solutions (2) imaginaires conjugués. On dit alors que G_0 est *stable*.

2°. Ou bien α est réel et différent de zéro et les deux solutions (2) réelles. On dit alors que G_0 est *instable*.

3°. Ou bien on a $\alpha = \alpha' + i\pi/U$, α' étant réel, et U étant la longueur totale de G_0; nos deux solutions peuvent alors se mettre sous la forme $v = e^{\alpha' u}\phi_1(u)$, $v = e^{-\alpha' u}\psi_1(u)$, les fonctions ϕ_1 et ψ_1 étant réelles et telles que

$$\phi_1(u + U) = -\phi_1(u), \qquad \psi_1(u + U) = -\psi_1(u).$$

On dit encore dans ce cas que G_0 est *instable*.

4°. Ou bien α est multiple de $i\pi/U$; c'est un cas limite, qui ne se présentera pas en général et que nous laisserons de côté pour le moment; dans ce cas les expressions (2) peuvent prendre une forme dégénérescente où figure $\log u$; ainsi qu'il arrive dans le cas des racines multiples pour les équations linéaires à coefficients constants.

Plaçons nous d'abord dans le 1er cas, et soient ρ et θ le module et l'argument de la 1ère solution (2); alors l'equation (1) admettra comme solutions particulières,

$$v = \rho e^{i\theta}, \qquad v = \rho e^{-i\theta},$$

et comme solution générale

$$v = A\rho \sin(\theta - \beta), \tag{3}$$

où β et A sont des constantes d'intégration.

Soit en particulier

$$v_1 = \rho \sin\theta, \qquad v_2 = \rho \cos\theta,$$

on aura

$$\operatorname{tg}\theta = \frac{v_1}{v_2} = B\int \frac{du}{v_2^2}, \tag{4}$$

B étant une constante. La formule (4) où tout est d'ailleurs réel nous montre que tg θ varie toujours dans le même sens, par exemple toujours en croissant; car si tg θ allait toujours en décroissant, il nous suffirait d'intervertir le rôle des deux solutions (2). Donc θ est une fonction croissante de u.

Supposons qu'on ait pris l'unité de longueur de telle façon que la longueur totale de G_0 soit égale à 2π. Qu'arrivera-t-il quand u augmentera de 2π? Il arrivera que ρ ne changera pas et que θ augmentera de $2\alpha\pi/i$ (ce qui nous montre d'ailleurs que α/i doit être positif si l'on veut que θ soit croissant).

Cela posé, nous appellerons *argument réduit* d'un point sur la géodésique fermée G_0 la quantité

$$\frac{i\theta}{\alpha}.$$

Si nous avons une unité de longueur quelconque et que U soit la longueur totale de G_0, l'argument réduit sera

$$\frac{2i\pi\theta}{\alpha U}.$$

Il résulte de là que l'argument réduit augmente constamment de 0 à 2π quand on parcourt la géodésique tout entière. Cet argument réduit peut donc servir à définir la position d'un point sur G_0.

Revenons à la solution générale (3); elle représente une géodésique G très peu différente de G_0; les points d'intersection successifs de G et G_0 seront donnés par la formule

$$\theta = \beta + K\pi,$$

où K est un entier quelconque.

Le $K^{\text{ème}}$ foyer du point $\theta = \beta$, sera donc le point $\theta = \beta + K\pi$; d'où cette conséquence: *Les arguments réduits d'un point et de son $K^{\text{ème}}$ foyer diffèrent d'une quantité constante et positive*

$$\frac{Ki\pi}{\alpha}\frac{2\pi}{U}.$$

L'analyse qui précède ne diffère pas de celle du No. 347 du tome 3 de mon ouvrage sur les méthodes nouvelles de la mécanique céleste. Au No. 349 j'ai envisagé les solutions instables et j'ai montré qu'elles se répartissent en trois catégories.

1°. Ou bien α est réel et $\phi(u)$ ne s'annule jamais (solutions instables de la 1ère catégorie). On a alors

$$e^{-2\alpha u}\frac{\psi(u)}{\phi(u)} = B\int \frac{du}{e^{-2\alpha u}\phi^2(u)},$$

ce qui montre que le 1er membre qui est toujours fini et croissant ne peut s'annuler qu'une fois; comme $e^{-2\alpha u}$ ne s'annule pas, il en résulte que $\psi(u)$ ne peut s'annuler qu'une fois; et comme $\psi(u)$ est périodique, $\psi(u)$ ne s'annule jamais. D'ailleurs la solution générale de (1) sera encore de le forme

$$v = e^{\alpha u}\phi(u)B\int \frac{du}{e^{-2\alpha u}\phi^2(u)}, \tag{5}$$

ce qui montre que v ne peut s'annuler qu'une fois: *Aucun point de G_0 n'a donc de foyer.*

2°. Ou bien α est réel et $\phi(u)$ s'annule (solutions instable, de la 2de catégorie).

La formule (5) nous montre alors que (l'intégrale du 2d membre étant toujours croissante et devenant infini quand ϕ s'annule) entre deux zéros consécutifs de $\phi(u)$, il y a un zéro de v et un seul.

En particulier entre deux zéros de $\phi(u)$, il y a un zéro de $\psi(u)$ et un seul. Si donc u_0, u_1, u_2, etc. sont les zéros successifs de $\phi(u)$, les valeurs de $\psi(u_0)$, $\psi(u_1)$, $\psi(u_2)$, $\cdots$ sont alternativement positifs et négatifs; mais la fonction $\psi(u)$ étant périodique, doit revenir à la même valeur quand on fait tout le tour de G_0. Donc *le nombre des zéros de $\phi(u)$ est toujours pair.* Soient donc

$$u_0, u_1, u_2, \cdots, u_n,$$

les zéros successifs de $\phi(u)$ (pair-n). Tout point situé entre u_0 et u_1 aura son foyer entre u_1 et u_2 et quand l'argument d'un point croîtra de u_0 à u_1 l'argument de son foyer croîtra *constamment* de u_1 à u_2; car s'il cessait de croître, c'est qu'un même point aurait plusieurs foyers, ce qui est impossible.

Supposons que α soit positif; alors pour u positif et très grand v sera sensiblement proportionnel à $e^{\alpha u}\phi(u)$; pour u négatif et très grand v sera sensiblement proportionnel à $e^{-\alpha u}\psi(u)$.

Les foyers d'ordre positif et très grand seront donc très voisins de l'un des zéros de $\phi(u)$.

Les foyers d'ordre négatif et très grand seront très voisins de l'un des zéros de $\psi(u)$. Soient alors

$$u_0', u_1', \cdots, u_n'$$

les zéros successifs de $\psi(u)$; u_0' étant compris entre u_0 et u_1.

Soient M_i et M'_i des points de G_0 qui correspondent respectivement à u_i et u'_i. Soit F_0 un point quelconque situé entre M_0 et M'_0, F_k son $K^{\text{ème}}$ foyer. Alors pour $k < n$, F_k sera compris entre M_k et M'_k; on voit ensuite que F_n est compris entre M_0 et M'_0; je dis que c'est entre M_0 et F_0; si en effet F_n était compris entre M'_0 et F_0, F_{2n} serait compris entre M'_0 et F_n et par conséquent entre M'_0 et F_0; plus généralement F_{kn} serait compris entre M'_0 et F_0. Donc F_{kn} ne pourrait tendre vers l'un des zéros de $\phi(u)$ quand k tend vers $+\infty$. En conséquence les points

$$F_n, F_{2n}, \cdots, F_{kn}, \cdots$$

sont entre M_0 et F_0 et quand K croît constamment et indéfiniment, ils se rapprochent constamment et indéfiniment de M_0. Au contraire les points

$$F_{-n}, F_{-2n}, \cdots, F_{-kn}$$

sont entre M'_0 et F_0 et quand K croît constamment et indéfiniment, ils se rapprochent constamment et indéfiniment de M'_0. Les zéros de $\phi(u)$ et ceux de $\psi(u)$ peuvent en conséquence recevoir le nom de *foyers limites*. Telle est la loi de la distribution des foyers pour les géodésiques fermées instables de la 2^{de} catégorie.

3°. Ou bien enfin α est de la forme $\alpha' + i\pi/U$, α étant réel (solutions instables de la $3^{\text{ème}}$ catégorie).

D'après ce que nous avons vu au No. 359 (loc. cit.), ces solutions jouissent des mêmes propriétés que celles de la 2^{de} catégorie. On a d'après ce qu'on a vu plus haut deux solutions de la forme:

$$v = e^{\alpha' u}\phi_1(u), \qquad v = e^{-\alpha' u}\psi_1(u),$$

avec

$$\phi_1(u+U) = -\phi_1(u), \qquad \psi_1(u+U) = -\psi_1(u),$$

et la solution générale s'écrit:

$$v = e^{\alpha' u}\phi_1(u)\, B \int \frac{du}{e^{-2\alpha' u}\phi_1^2(u)}. \tag{5 bis}$$

Nous voyons que $\phi_1(u)$ doit s'annuler, puisque cette fonction change de signe, qu'entre deux zéros de $\phi_1(u)$, il y a un zéro de $\psi_1(u)$ et un seul.

Ce que nous avons dit de la distribution des foyers successifs et de leurs relations avec les zéros de ϕ et de ψ (dont le rôle est joué ici par les zéros de ϕ_1 et de ψ_1) subsiste sans changement: la seule différence c'est que *le nombre des zéros de* $\phi_1(u)$ *au lieu d'être toujours pair, est toujours impair.*

En effet la fonction $\psi_1(u)$ doit, non pas revenir à la même valeur, mais au contraire changer de signe quand on fait tout le tour de G_0.

Dans le cas des surfaces convexes tout point a un foyer dont la distance à ce point est inférieure à une limite donnée; il n'y a donc pas de géodésiques fermées instables de la $1^{\text{ère}}$ catégorie; il y en a au contraire pour les surfaces à courbures opposées.

Au No. 355 (loc. cit.) j'ai montré que la condition nécessaire et suffisante pour qu'une trajectoire fermée corresponde à un minimum de l'action, c'est qu'elle soit instable de la 1ère catégorie. Il résulte de là que sur une surface convexe, il n'y a pas de géodésique fermée qui soit plus courte que toutes les courbes fermées infiniment voisines. Cela est d'ailleurs aisé à établir directement.

Considérons un ellipsoïde et ses trois ellipses principales; la plus grande et la plus petite sont des géodésiques fermées stables; la moyenne est une géodésique fermée instable de la 2de catégorie. Sur un hyperboloïde à 1 nappe, l'ellipse de gorge est une géodésique fermée instable de la 1ère catégorie.

Il est aisé de voir (loc. cit., No. 378) que si deux géodésiques fermées viennent à se confondre, pour disparaitre ensuite en devenant imaginaires (lorsqu'on fait varier d'une manière continue le paramètre que nous appelions t dans le § précédent), l'une d'elles est stable et l'autre instable. D'où il resulte que dans une même série continue, l'excès du nombre des géodésiques stables sur celui des géodésiques instables est constant. Or nous avons vu que les géodésiques fermées sans point double forment une ou plusieurs séries continues, et d'autre part que l'on peut toujours passer d'une manière continue d'une surface convexe quelconque à une autre.

Donc *si sur une surface convexe quelconque on envisage toutes les géodésiques fermées sans point double l'excès du nombre de celles qui sont stables sur le nombre de celles qui sont instables est constant*; il est donc le même que pour l'ellipsoïde, *il est donc égal à* 1.

Sur une surface convexe, il y a donc toujours au moins une géodésique fermée stable sans point double.

Remarquons que α que l'on appelle *l'exposant caractéristique* de G_0 est par ce qui précède entièrement déterminé. On pourrait croire le contraire, car si U est la longueur totale de G_0, on a

$$e^{\alpha u}\phi(u) = e^{\alpha + \frac{2i\pi u}{U}}\left[\phi(u)e^{-\frac{2i\pi u}{U}}\right],$$

où $\phi(u)e^{-2i\pi u/U}$ est comme $\phi(u)$ une fonction périodique. Il semble donc que l'on puisse remplacer α par $\alpha + 2i\pi/U$ ou par $\alpha + 2ki\pi/U$ (k entier positif ou négatif). Mais dans le cas des géodésiques instables, il n'y a qu'une des quantités $\alpha + 2ki\pi/U$ qui soit réelle et c'est celle qu'il convient d'adopter.

Dans le cas des géodésique stables, α est entièrement déterminé par la condition que l'argument θ augmente de $\alpha U/i$ quand on fait le tour de la géodésique G_0. Autre remarque: considérons une géodésique instable de la 2de ou de la 3ème catégories; elle fera partie d'une série continue au sens du § 4. Je dis que le nombre des zéros de $\phi(u)$, d'où dépend comme nous venons de le voir, la loi de distribution des foyers, demeurera constant; car il ne pourrait varier que si $\phi(u)$ et $\phi'(u)$ s'annulaient en même temps, ce qui ne peut avoir lieu. Une

série continue peut se partager en *sections*, de telle façon que dans une section par exemple toutes les géodésiques soient stables, dans la suivante toutes instables, dans la suivante toutes stables, et ainsi de suite. Aux points où l'on passe d'une section à l'autre, l'exposant caractéristique est multiple de $i\pi/U$. Eh bien *dans une même section d'une même série continue, le nombre des foyers limites est invariable.*

Dans le cas des géodésiques instables, il y a plusieurs manières de définir l'exposant α; nous pouvons, comme nous l'avons fait jusqu'ici, supposer α réel pour celles de la 1ère et de la 2de catégorie, tandis que pour celles de la 3ème catégorie, on supposera la partie imaginaire de α égale à $i\pi/U$. Mais on peut également adopter une autre solution.

On peut supposer que la partie imaginaire de α est égale à $ni\pi/U$, n étant le nombre des zéros de $\phi(u)$ (ou de ϕ_1); l'exemple qui va suivre nous fera comprendre les raisons qui pourraient justifier cette convention. Envisageons la surface

$$x^2 + y^2 + \frac{q^2 z^2}{4} - \frac{q_1 z^2 x}{4} = 1,$$

q et q_1 étant deux constantes.

Cette surface est convexe et fermée si q_1 est suffisamment petit; car elle diffère très peu de l'ellipsoïde de révolution

$$x^2 + y^2 + \frac{q^2 z^2}{4} = 1.$$

Elle admet une géodésique fermée qui est le cercle

$$z = 0, \qquad x^2 + y^2 = 1.$$

Si pour ce cercle, nous exprimons x et y en fonctions de l'arc u, nous trouvons

$$x = \cos u, \qquad y = \sin u.$$

L'équation des géodésiques infiniment peu différentes de ce cercle sera

$$v'' + \tfrac{1}{4}(q^2 - q_1 \cos u)v = 0.$$

Nous reconnaissons une équation qui a fait l'objet de travaux nombreux (loc. cit., Tome 2, Chapitre 17), les notations sont les mêmes que dans ce chapitre 17 à la condition de changer x en v et $2t$ en u. Le rôle de l'exposant α est le même que celui que jouait dans l'ouvrage cité la quantité

$$\frac{ih}{2}.$$

Si donc q_1 est petit nous avons sensiblement

$$\alpha = \frac{iq}{2}.$$

D'ailleurs α est imaginaire et la géodésique stable, à moins que q ne soit voisin d'un entier.

Si q est voisin d'un entier, n de telle façon que sur la figure de la page 243 (loc. cit.) le point de coordonnées q, q_1 se trouve dans une des régions couvertes de hachures, la géodésique est instable. D'après le No. 352 (loc. cit., Tome 3) elle est toujours de la 2de ou de la 3ème catégorie).

Quel est le nombre des zéros de $\phi(u)$ ou de $\phi_1(u)$? Il suffit de remarquer que sur l'une des courbes qui limitent les régions couvertes de hachures (loc. cit., No. 352) la fonction ϕ ou ϕ_1 est impaire et se réduit sensiblement (à cause de la petitesse de q_1) à sin $qu/2$, ou à sin $nu/2$ puisque q est très voisin de n.

Le nombre des zéros de $\phi(u)$ ou de $\phi_1(u)$ est donc n et celui des foyers limites est $2n$, dont n pour les zéros de $\phi(u)$ et n pour les zéros de $\psi(u)$. Cela nous montre:

1°. qu'il existe des surfaces qui admettent des géodésiques fermées instables de la 2de et de la 3ème catégories, le nombre des foyers limites pouvant être quelconque.

2°. que si l'on veut que α soit une fonction continue, il faut lui attribuer pour partie imaginaire $ni\pi/U$, $2n$ étant le nombre des foyers limites.

§ 6. *De quelques types de géodésiques fermées.*

Soit G_0 une géodésique fermée stable dont la longueur totale est U et qui appartient à une série continue. Soit α son exposant caractéristique. Il est clair que U et α sont des fonctions continues du paramètre que nous avons appelé t au § 1 et qui correspond à cette série continue.

Supposons que pour une certaine valeur de t, pour $t = t_0$ par exemple, le produit αU soit commensurable avec $i\pi$ par, exemple égal à

$$2\frac{mi\pi}{n},$$

m et n étant deux entiers premiers entre eux. L'argument réduit d'un point et de son $K^{\text{ième}}$ foyer différeront de

$$\pi\frac{nK}{m},$$

ce qui montre que tout point de notre géodésique G_0 coïncidera avec son $2m^{\text{ième}}$ foyer. Si l'on considère un point M de G_0, et la géodésique issue du point M et infiniment voisine de G_0 on voit qu'elle viendra repasser en M, aux infiniment petits près d'ordre supérieur après avoir recoupé G_0 $2m$ fois (le point M compris). Elle aura ainsi fait n fois le tour de la géodésique G_0. Les principes exposés dans le Chapitre 28 (loc. cit.) nous permettent d'interpréter ce résultat. Il existe outre la géodésique G_0, deux autres géodésiques fermées G_1 et G_1' jouissant des propriétés suivantes:

1°. Pour $t < t_0$ (par exemple) ces deux géodésiques n'existent pas ou plutôt sont imaginaires; 2°. Pour $t > t_0$ elles sont toutes deux réelles; 3°. Pour $t = t_0$ elles se confondent entre elles et se réduisent l'une et l'autre à la géodésique G_0 parcourue n fois; 4°. Elles appartiennent à une même série continue; 5°. l'une d'elles est stable et l'autre instable.

La relation de G_1 et G_1' avec G_0, est la même que celle qu'on rencontre en mécanique céleste, entre les solutions périodiques de la 2de et de la 3ème sorte et celles de la 1ère sorte, ou entre les solutions du 2d genre et celles du 1er genre. La série continue S_1 à la quelle appartiennent G_1 et G_1' peut donc être regardée comme *engendrée* par la série S_0 à la quelle appartient G_0; comme le produit αU, en variant d'une manière continue, passera une infinité de fois par des valeurs commensurables avec $i\pi$, la série S_0 engendrera une infinité d'autres séries continues. Et comme l'une des deux géodésiques G_1 et G_1' est stable, le série S_1 va à son tour engendrer de la même manière une infinité d'autres séries.

Nous avons vu qu'il y a toujours une géodésique fermée stable, sans point double. Appelons-la G_0. Supposons qu'elle engendre comme nous venons de l'expliquer deux autres géodésiques fermées G_1 et G_1'; soit $mi\pi/n$ la valeur correspondante de αU. Pour des valeurs de t très voisines de t_0, ces deux géodésiques différeront très-peu de G_0 et nous obtiendrons leur équation approximative en reprenant l'analyse du § précédent; nous trouverons alors (en reprenant les équations (3) du § 2 précédent)

$$v = A\rho \sin(\theta - \beta) \quad \text{pour } G_1,$$

$$v = A'\rho \sin(\theta - \beta') \quad \text{pour } G_1',$$

A, β, A', β' étant des constantes.

Les points d'intersection de G_0 et de G_1 par exemple nous sont donnés par la formule

$$\theta = \beta + K\pi \quad (K \text{ entier}),$$

et les valeurs correspondantes de l'argument réduit sont

$$\frac{2i\pi\beta}{\alpha U} + \frac{K\pi \cdot 2i\pi}{\alpha U} = \frac{n\beta}{m} + \frac{n}{m} K\pi.$$

Combien l'expression

$$\frac{nK}{m},$$

peut-elle prendre de valeurs distinctes, deux valeurs n'étant pas regardées comme distinctes lorsqu'elles diffèrent d'un entier pair? Evidemment $2m$; donc *le nombre des points d'intersection de G_0 avec G_1, et par conséquent avec une géodésique quelconque de le série S_1 est égal à $2m$.*

Cherchons maintenant le nombre des points doubles. Supposons la géodésique fermée G_1 partagée en $2m$ arcs par ses $2m$ points d'intersection avec G_0; ces

arcs seront alternativement d'un côté et de l'autre de G_0; d'où il suit qu'un arc de rang pair ne peut pas couper un arc de rang impair, mais que deux arcs de rang pair, ou deux arcs de rang impair peuvent se couper.

J'observe ensuite que $v =$ partie réelle $e^{au}\phi(u)$ est une fonction uniforme de u de sorte que les points doubles correspondront aux valeurs de u pour lesquelles on aura

$$(h) \qquad v(u) = v(u + 2K\pi),$$

K étant entier, et qu'il n'y en aura pas d'autre. Nous avons d'ailleurs

$$\frac{v(u + 2K\pi)}{v(u)} = B\int \frac{du}{v^2(u)},$$

car $v(u)$ et $v(u + 2K\pi)$ sont deux intégrales de l'équation linéaire (1) du § précédent ce qui montre que le rapport de ces deux intégrales va constamment en croissent ou constamment en décroissant. Entre deux zéros consécutifs de $v(u)$, qui correspondent aux valeurs $+\infty$ et $-\infty$ de ce rapport, il prendra donc une fois et une seule la valeur zéro, une fois et une seule la valeur 1. Entre deux points d'intersection consécutifs de G_0 et de la branche $v(u)$ de la géodésique G_1 (points donnés par l'équation $v(u) = 0$), il y aura un point d'intersection de G_0 et de la branche $v(u + 2K\pi)$ de la géodésique G_1 (point donné par l'équation $v(u + 2K\pi) = 0$) et il n'y en aura qu'un seul; de plus il y aura un point double donné par l'équation (4) et il n'y en aura qu'un seul.

Cela nous permet d'énumérer les points doubles. Donnons à l'entier K une valeur déterminée et faisons varier u de 0 à $2n\pi$, nous aurons $2m$ points d'intersection de G_0 avec la branche $v(u)$ de G_1; cela nous donnera donc $2m$ points doubles. Maintenant nous pouvons donner à K les valeurs $1, 2, \cdots n-1$, ce qui fait en tout $2m(n-1)$ points doubles. Mais chaque point double est ainsi compté deux fois, puisqu'on retrouve le même en changeant u et K en $u + 2K\pi$ et $n - K$.

Le nombre total des points doubles pour la géodésique G_1 qui diffère infiniment peu de G_0, et par conséquent aussi pour toutes les géodésiques fermées de la série S_1, est donc $m(n-1)$.

Pour $n = 1$, on retrouve des géodésiques fermées sans point double; pour $n = 2$ on constate la circonstance suivante: Les divers arcs de G_1 partagent la surface en $m + 2$ regions, dont m sont des polygones curvilignes de deux côtés et les deux autres des polygones curvilignes de m côtés. Si m n'est pas égal à 3 aucun de ces polygones n'est un triangle, de sorte que d'après ce que nous avons dit au § 4, la disposition des points doubles ne peut varier quand on parcourt la série S_1 d'une manière continue; *toutes les géodésiques de cette série appartiennent donc au même type.*

Il semble qu'il y ait exception pour $m = 3$ et que deux types soient possibles,

représentés par les figures 1 et 2; mais la disposition de la figure 2 est impossible. En effet la courbe 2 limite quatre régions A, B, C, D; la région D est un triangle dont les trois angles sont α, β, γ; la courbure totale de D est $\alpha+\beta+\gamma-\pi$; de même celles de A, B, C sont respectivement $\alpha+\pi$, $\beta+\pi$, $\gamma+\pi$; de sorte que

$$A+B+C>D+4\pi>4\pi,$$

chacune des grandes lettres représentant la courbure totale de la région correspondante. Ce qui voudrait dire que la courbure totale des trois régions

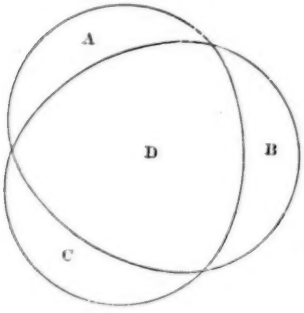


FIG. 1.

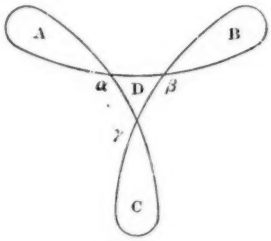


FIG. 2.

A, B, C réunies serait plus grande que celle de la surface entière, résultat évidemment absurde.

Supposons maintenant une géodésique G_1 de la série S_1 qui ne soit plus très-peu différente de G_0. Je considère le produit αU correspondant et je suppose qu'il passe par la valeur

$$\frac{2m_1 i\pi}{n_1},$$

commensurable avec $i\pi$. D'après ce que nous venons de voir, la série S_1 engendrera une autre série S_2; soit G_2 une des géodésiques de cette série et par exemple celle qui diffère très peu de la géodésique G_1 parcourue n_1 fois. On voit d'abord qu'elle rencontrera G_0 en $2mn_1$ points. En combien de points rencontre-t-elle G_1? Les points d'intersection seront de deux sortes. D'abord à chaque point double de G_1 correspondront $2n_1$ points d'intersection. En effet en chaque point double viennent passer 2 branches de G_1, soient B et B', et alors B rencontre n_1 branches de G_2 très peu différentes de B'; et B' rencontre n_1 branches de G_2 très peu différentes de B. On obtient ainsi $2mn_1(n-1)$ points d'intersection. Il faut y ajouter les $2m_1$ points d'intersection de G_2 avec la branche correspondante de G_1, soit en tout

$$2mn_1(n-1)+2m_1.$$

Combien G_2 aura-t-elle de points doubles? Chaque point double de G_1 nous

donnera d'abord n_1^2 points doubles de G_2, car nous avons les n_1 branches peu différentes de B, qui coupent en n_1^2 points, les n_1 branches peu différentes de B'; nous avons en outre $m_1(n_1-1)$ points doubles obtenus par des équations analogues à (4), soit en tout

$$mn_1^2(n-1)+m_1(n_1-1).$$

Ainsi G_2 et par conséquent toutes les géodésiques de la série S_2 couperont G_0 en $2mn_1$ points, G_1 en $2mn_1(n-1)+2m_1$ points et auront $mn_1^2(n-1)+m_1(n-1)$ points doubles.

Il y a une autre quantité qui demeure invariable dans toutes les géodésiques d'une même série, c'est la courbure totale de la région limitée par la géodésique; dans le cas d'une courbe fermée sans point double, c'est 2π; dans le cas de la figure 1, c'est 4π en comptant A, B et C une fois, et D deux fois, conformément aux conventions habituelles.

§ 7. *Existence d'une géodésique fermée.*

Nous avons vu plus haut qu'il y a toujours au moins une géodésique fermée sans point double. Bien que la démonstration ne laisse rien à désirer au point de vue de la simplicité, je crois cependant devoir en donner une seconde, quoique beaucoup moins simple.

Soit une géodésique fermée sans point double, elle partagera la surface en deux régions et la courbure totale de chacune de ces deux régions sera 2π. Considérons maintenant toutes les courbes fermées sans point double qui partagent la surface en deux régions dont la courbure totale est 2π. La longueur de l'une de ces courbes fermées ne peut pas devenir plus petite que toute quantité donnée; car si cette courbe se réduisait à un contour infiniment petit, la courbure totale de l'une des régions limitées par ce contour serait elle-même infiniment petite.

Parmi ces courbes, il y en a donc une qui est plus courte que toutes les autres, et je dis que c'est une géodésique.

Contentons-nous d'abord d'un premier aperçu, afin de faire compendre le principe de la démonstration.

Soit C une courbe fermée quelconque et R l'une des régions limitées par C, soit

$$U=\int ds, \qquad \Omega=\int \frac{d\omega}{\rho},$$

la longueur totale de C et la courbure totale de R; dans ces formules ds représente l'élément d'arc de C, $d\omega$ l'élément de surface de R, et ρ le produit des deux rayons de courbure principaux.

Considérons une courbe C' très peu différente de C et limitant une région

R', soit ζ la distance de ces deux courbes comptée sur la normale à C. Soit $U + \delta U$ la longueur de C', et $\Omega + \delta\Omega$ la courbure totale de R' ; on aura

$$\delta U = \int \gamma\zeta ds\,, \qquad \delta\Omega = \int \frac{\zeta ds}{\rho}\,,$$

γ représentant la courbure géodésique de C. Nous nous sommes imposé la condition $\Omega = 2\pi$, et nous voulons que U soit minimum ; nous devons donc avoir

$$\delta U = 0\,, \qquad \delta\Omega = 0\,.$$

Ces deux équations doivent donc être équivalentes, c'est à dire qu'on doit avoir

$$\gamma = \frac{K}{\rho}\,,$$

K étant une constante.

Mais pour une courbe fermée quelconque C limitant une région R on a

$$\int \gamma ds = \Omega - 2\pi\,.$$

Dans notre cas, on a $\Omega = 2\pi$; d'où

$$K\int \frac{ds}{\rho} = 0\,.$$

Comme la surface est convexe, ρ est essentiellement positif de sorte que l'intégrale $\int ds/\rho$ ne peut s'annuler ; on aura donc

$$K = 0\,, \qquad \gamma = 0\,,$$

ce qui veut dire que *la courbe C est une géodésique.*

Examinons maintenant les objections qu'on pourrait faire à ce raisonnement incomplet. Considérons l'ensemble des courbes analytiques C fermées et sans point double, limitant une région R de courbure totale 2π ; il est clair que la longueur de ces courbes aura une limite inférieure ; mais on peut se demander si cette limite sera effectivement atteinte ; et si elle le sera par une courbe faisant partie de l'ensemble.

Cette courbe pour laquelle le minimum serait atteint sera-t-elle analytique, de telle façon que les règles du calcul des variations puissent lui être appliquées ?

Et si elle est analytique, sera-t-elle dépourvue de points doubles, ou bien ne pourrait-il se faire que tout en étant infiniment voisine de courbes sans point double elle possédât elle-même des points singuliers où deux de ses branches viendraient à se toucher ?

Cela fait en réalité deux objections distinctes.

La 1ère est d'ordre analytique ; c'est celle que l'on rencontre dans tous les problèmes analogues ; malgré son importance, je n'en parlerai pas ici, je me bornerai à renvoyer aux récents travaux de M. HILBERT.

La 2^de^ est d'ordre physique, pour ainsi dire ; elle est spéciale au problème qui nous occupe et nous allons en faire un examen approfondi.

Pour la bien faire comprendre, nous allons donner à notre problème une signification physique et concrète. Supposons d'abord un gaz enfermé dans une enceinte en partie déformable, il exercera sur les parois de cette enceinte une pression qui tendra à en augmenter le volume, le travail effectué par la pression de ce gaz dépendra des variations de ce volume. Au lieu d'un gaz, nous pouvons supposer un fluide compressible comme un gaz, mais tel que la pression soit liée au volume non par la loi de Mariotte, mais par une relation quelconque. Alors le travail virtuel de la pression pour une déformation virtuelle quelconque de l'enceinte sera $p\delta V$, p désignant la pression du gaz, laquelle doit être uniforme, et δV l'accroissement virtuel du volume.

Si l'enceinte n'est susceptible que de certaines déformations, l'equilibre sera atteint quand le volume sera maximum, en supposant que toutes les déformations possibles se fassent sans résistance. Si au contraire certaines forces opposent à la déformation de l'enceinte et que le travail virtuel de ces forces soit δW la condition d'équilibre sera

$$p\delta V + \delta W = 0.$$

Supposons maintenant que notre enceinte soit limitée : 1°. Par notre surface convexe S elle-même, 2°. Par une autre surface convexe S' infiniment peu différente de S et telle que la distance des deux surfaces, comptée suivant la normale à S, soit égale à ϵ/ρ, ϵ étant une constante infiniment petite et ρ le produit des rayons de courbure, 3°. Par un ruban infiniment étroit dont les bords s'appuieront sur S et sur S', (c'est ce ruban qui va épouser la forme de notre courbe fermée C, de telle façon que notre enceinte sera l'espace infiniment mince compris entre la région de S que nous avons appelée R et la région correspondante de S'). Le volume de notre enceinte sera alors représenté par l'intégrale

$$\int \frac{\epsilon}{\rho} d\omega = \epsilon\Omega.$$

Nous supposerons que notre ruban résiste plus on moins à l'extension, mais qu'il est d'ailleurs flexible sans résistance ; sa largeur pourra être supposée une peu plus grande que la plus grande valeur de ϵ/ρ, de telle sorte que ses bords s'appliquent sur les surfaces S et S' et soient *collés* contre ces surfaces par la pression du fluide. La pression du fluide tendra à allonger le ruban et mettra en jeu la résistance du ruban à l'extension, c'est à dire ce qu'on appelle la *tension* du ruban ; cette tension sera constante tout le long du ruban, puisque la pression du fluide est normale à ce ruban ; nous l'appellerons T ; on aura alors

$$\delta W = - T\delta U,$$

U étant la longueur totale du ruban, c'est à dire la courbe C. L'équation d'équilibre s'écrit alors

$$p\epsilon\delta\Omega = T\delta U.$$

Si le fil est inextensible, c'est à dire si un accroissement très petit de U amène un accroissement très grand de S: la longueur U est constante et doit être regardée comme une des données de la question, et l'équilibre est atteint quand Ω est maximum.

Si au contraire le fluide est incompressible, c'est à dire si une diminution très petite du volume amène un accroissement tres grand de p, c'est Ω qui est constant et qui est une donnée de la question (comme dans le cas qui nous occupe où $\Omega = 2\pi$) et l'équilibre est atteint quand U est minimum.

Considérons un segment très petit ds du ruban, limité en AB et $A'B'$; tout se passera comme si le ruban était coupé en AB et $A'B'$ et soumis à deux forces appliquées l'une au milieu de AB, l'autre au milieu de $A'B'$ et représentant l'une l'action de la portion du ruban située au dela de AB, l'autre l'action de la portion du ruban située au dela de $A'B'$; ces deux forces ne sont autre chose que la tension du ruban; elles sont donc égales en grandeur et égales à T; elles doivent faire équilibre à la pression exercée par le fluide sur le segment $ABA'B'$. Projetons tout sur la normale à la courbe C, dans le plan tangent à la surface S. La projection de la pression du fluide sera

$$\frac{p\epsilon}{\rho}ds,$$

celle de la tension sera $Td\alpha$, $d\alpha$ étant l'angle de contingence géodésique de C. On aura donc

$$\frac{p\epsilon}{\rho}ds = Td\alpha.$$

Comme T, p et ϵ sont des constantes, nous voyons que le rapport de la courbure géodésique $d\alpha/ds$ à $1/\rho$ doit être une constante; nous retrouvons ainsi, dans un autre langage, le résultat obtenu plus haut.

L'équilibre pourra certainement être atteint d'une façon quelconque, ce qui ne peut se faire que quand C sera une géodésique fermée *à moins que*, et c'est ici que nous retrouvons la difficulté signalée plus haut, deux portions du ruban ne viennent se coller l'une contre l'autre. Dans ce cas les deux portions du ruban ainsi collées l'une sur l'autre, subiraient la pression du fluide des deux côtés, de telle façon que l'effet de cette pression se trouverait annulée.

Dans ce cas l'équation d'équilibre s'écrirait, pour cette portion du ruban,

$$Td\alpha = 0,$$

de sorte que l'angle de contingence géodésique étant nul cette partie du ruban devrait affecter la forme d'une géodésique.

Ainsi nous devrons distinguer les *parties libres* du ruban qui devraient satisfaire à la condition $\rho d\alpha/ds = \text{const.}$, et les *parties collées* qui se réduiraient à des arcs de géodésiques. Que se passerait-il maintenant aux points de raccordement? Je dis que les diverses parties du ruban devraient se raccorder par contact, c'est à dire sous un angle nul et de façon que l'ensemble du ruban ne présente pas de point anguleux.

Soient en effet BAC, EDF deux brins du ruban, AB étant appliqué sur ED; nous savons que la tension est constante tout le long de AB, de même que tout le long de AC, de ED ou de DF; mais nous ne savons pas encore si elle est la même par exemple sur AC et sur AB.

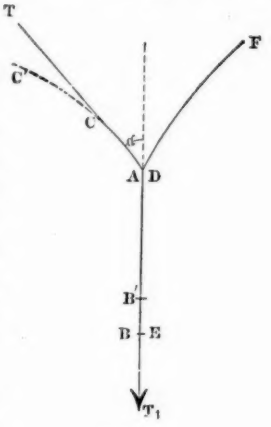


FIG. 3.

Mais il est aisé d'établir ce dernier point; le brin BAC est en équilibre sous l'action des deux tensions CT et ET' appliquées à ses deux extrémités, de la pression du fluide et de l'action exercée par le brin DE sur le brin AB sur lequel il s'applique. Cette action (égale d'ailleurs au sens près à la réaction du brin AB sur le brin DE) est normale au brin AB, puisque les deux brins, quoique appliqués l'un sur l'autre, peuvent librement glisser l'un sur l'autre.

Supposons alors que le brin BAC se déplace pour venir en $B'AC'$ (c'est à dire glisse en passant par le point fixe A); la somme des travaux virtuels doit être nulle. Or le travail de la pression et celui de l'action de DE sur AB sera nul, puisque ces forces sont normales aux brins sur lesquels elles s'exercent. Les travaux des deux tensions seront donc égaux et de signe contraire, ce qui exige que les deux tensions CT et BT, soient égales. C. Q. F. D.

Supposons maintenant les arcs AC et AB très petits; soit AM le prolongement de la tangente en A à AB; soit α l'angle de CT avec AM. Projetons sur une perpendiculaire à AM toutes les forces qui aggissent sur ABC. Les pressions seront négligeables puisque les arcs sont très petits de sorte que le brin ABC devra être en équilibre sous l'action des deux tensions et de l'action de DE sur AB; il faut donc (puisque les tensions sont égales) que cette action soit dirigée suivant la bissectrice des deux tensions; c'est à dire qu'elle fasse avec la perpendiculaire à AB un angle $\alpha/2$; or elle est normale à AB. Donc l'angle de raccordement α est nul. C. Q. F. D.

Ainsi le ruban forme une courbe sans point anguleux le long de laquelle la tension T est constante.

Alors la courbure totale de R sera encore

$$\frac{p\epsilon}{T}\int\frac{ds}{\rho} + 2\pi,$$

l'intégrale étant étendue à toute la partie libre du ruban; on aura donc

$$\frac{p\epsilon}{T}\int\frac{ds}{\rho}=0,$$

ce qui veut dire encore que $p=0$, ou que la courbure géodésique est partout nulle, ou que le ruban entier affecte partout la forme d'une géodésique fermée.

Si alors AB et A_1B_1 sont deux parties du ruban collées l'une sur l'autre, si AC et A_1C_1 sont les parties libres du ruban au delà de A et de A_1; si BD et B_1D_1 sont les parties libres du ruban au delà de B et de B_1, toutes ces parties devront appartenir à une même géodésique; donc AC est la continuation analytique de AB, et A_1C_1 celle de A_1B_1. Ce qui est absurde, puisque par hypothèse AB coïncide avec A_1B_1 et que AC ne coïncide pas avec A_1C_1. L'hypothèse où deux parties du ruban viennent se coller l'une sur l'autre doit donc être écartée.

Nous devons de même écarter celle où deux parties du ruban viendraient se toucher en un point, car nous avons vu que deux géodésiques ne peuvent être tangentes l'une à l'autre.

L'équilibre ne pourra donc être atteint que quand le ruban prendra la forme d'une géodésique fermée sans point double.

§ 8. *Discussion du minimum.*

La courbe C dont la longueur est un minimum, doit d'après ce qui précède se réduire à une géodésique; mais cette condition nécessaire n'est pas suffisante; c'est celle qui se rapporte à la 1ère variation et nous avons à examiner celle qui se rapporte à la 2de variation.

Soit C la courbe considérée, qui est représentée sur la figure en $AMBNEPA$. Supposons qu'il existe une courbe AM_1BN_1E infiniment peu différente de la 1ère et satisfaisant comme elle à la condition:

$$\text{Courbure géodésique}=\frac{K}{\rho},$$

K étant une constante qui pourra ne pas avoir la même valeur que pour la courbe C (c'est à dire la valeur zéro).

Je suppose de plus que la courbure totale de l'aire infiniment petite AM_1BMA, soit égale à la courbure totale de l'aire $BNEN_1B$, de telle façon que les deux courbes fermées $AMBNEPA$ et AM_1BN_1EPA enveloppent des aires de même courbure totale.

Je considère les angles sous lesquels les deux courbes $AMBNE$, AM_1BN_1E se coupent en A_1 en B et en E comme des infiniment petits du 1er ordre, et je remarque d'abord que la différence des longueurs des deux courbes est infiniment petite du 3ème ordre.

Prenons sur AM_1B et AP deux points A'' et A' voisins de A sans en être infiniment voisins (je les suppose simplement assez près pour que les géodésiques qui les joignent en restant très voisins de C restent le plus court chemin d'un point à un autre); et de même sur EN_1B et EP deux points E'' et E' voisins de E. Joignons par des arcs de géodésique marqués en traits pointillés $A'A''$ d'une part, $E'E''$ d'autre part. Il est clair:

1°. Que la longueur $A'A''$ sera plus petite que $A'A + AA''$, et de même la longueur $E'E''$ plus petite que $E'E + EE''$, et que les différences seront des infiniment petits du 2d ordre (et non pas d'ordre supérieur au 2d); en effet

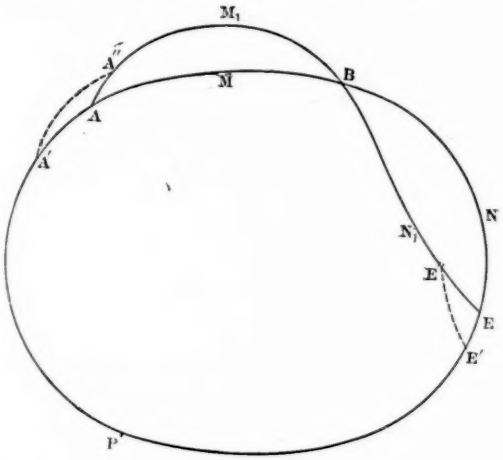


FIG. 4.

$A'A''$ étant une géodésique $AA' + AA'' - A'A''$ par exemple, est au moins égal à $\overline{AA'} + \overline{AA''} - A'A''$, en designant par $\overline{AA'}$ et $\overline{AA''}$ les géodésiques qui joignent A à A' ou à A''; (comme AA' est une géodésique, elle coïncide avec $\overline{AA'}$). Les deux géodésiques se coupent sous de angles infiniment petits du 1er ordre, (et non d'ordre supérieur, puisque la distance de A'' à C est du 1er ordre) de sorte que $\overline{AA'} + \overline{AA''} - A'A''$ est positif et plus grand qu'un infini petit du 2d ordre.

Il résulte de tout cela que la courbe fermée $A'A''M_1BN_1E''E'PA'$ est plus courte que la courbe fermée $AMBNEPA$, ou C.

2°. Il est clair ensuite que l'on peut placer les points A', A'', E', E'' de telle façon que les deux triangles $AA'A''$ et $EE'E''$ aient même courbure totale. Dans ces conditions, les deux courbes $AMBNEPA$ et $A'A''M_1BN_1E''E'PN$ enveloppent des aires ayant même courbure totale.

Si donc on peut construire une courbe AM_1BN_1E placée comme sur la figure et jouissant des propriétés énoncées, C n'est certainement pas la plus courte de

toutes les courbes fermées qui enveloppent des aires de courbure totale 2π. *La courbe C ne correspond pas à un vrai minimum.* Supposons maintenant que C soit une géodésique fermée stable et que le produit $\alpha U/i$ soit plus grand que 2π. Je dis que C ne pourra pas correspondre à un vrai minimum.

Soit en effet A un point quelconque de C, AM_1B une géodésique très voisine de C et passant par A; elle viendra recouper C en un point B qui sera le 1er foyer de A puis en un point E qui sera le 2d foyer de A. Si $\alpha U/i$ est $> 2\pi$, les deux arcs AMB, BNE ne couvriront pas le périmètre de C tout entier et les deux courbes seront disposées comme sur la figure.

Si de plus l'angle sous lequel les deux géodésiques se coupent en A est égal à celui sous lequel les deux géodésiques se coupent en E, les deux aires AM_1BA, BN_1EMB auront même courbure totale; nous serons dans les conditions du théorème précédent et C ne sera pas un vrai minimum.

La question à résoudre est donc la suivante: peut-on choisir le point A de telle façon que les deux angles en question soient égaux?

Reprenons les rotations du § 5, la géodésique AM_1B aura pour équation

$$v = \rho \sin(\theta - \beta),$$

et les points A, B, E correspondront à

$$\theta = \beta, \qquad \theta = \beta + \pi, \qquad \theta = \beta + 2\pi,$$

l'angle sous lequel les deux géodésiques se couperont en chacun de ces trois points sera représenté par la valeur correspondante de $\rho d\theta/du$. La condition à remplir est donc

$$\left(\rho \frac{d\theta}{du}\right)_{\theta=\beta} = \left(\rho \frac{d\theta}{du}\right)_{\theta=\beta+2\pi}.$$

Mais ρ et $d\theta/du$ sont des fonctions périodiques de u se reproduisant quand u augmente de U. Ce sont donc aussi des fonctions périodiques de l'argument réduit

$$\frac{2i\pi\theta}{\alpha U}.$$

Nous aurons donc

$$\rho \frac{d\theta}{du} = F\left(\frac{2i\pi\theta}{\alpha U}\right),$$

F étant développable en série de Fourier suivant les cosinus et les sinus des multiples de $2i\pi\theta/\alpha U$.

Il s'agit donc de savoir si l'on peut déterminer β de telle façon que

$$F\left(\frac{2i\pi\beta}{\alpha U} + \frac{4i\pi^2}{\alpha U}\right) - F\left(\frac{2i\pi\beta}{\alpha U}\right) = 0.$$

Or le premier membre est une fonction périodique de $2i\pi\beta/\alpha U$, développable en

série trigonométrique suivant les sinus et les cosinus des multiples de $2i\pi\beta/\alpha U$, *et dont la valeur moyenne est nulle.* Ce premier membre ne peut donc être toujours de même signe et il faut bien qu'il s'annule. *Donc C n'est pas un vrai minimum.* C. Q. F. D.

Supposons maintenant que C soit instable, soit n le nombre des zéros des $\phi(u)$ (ou de ϕ_1) et par conséquent n le nombre des zéros de $\psi(u)$ et $2n$ celui des foyers limites. Soient

$$u_0, u_1, u_2, \cdots, u_{n-1}, u_n = u_0 + U, u_{n+1} = u_1 + U,$$

les zéros successifs de $\phi(u)$ (ou de ϕ_1); soient

$$u'_0, u'_1, u'_2, \cdots, u'_{n-1}, u'_n = u'_0 + U, u'_{n+1} = u'_1 + U,$$

les zéros successifs de $\psi(u)$.

Les zéros de $\phi(u)$ partagent le périmètre de C en n segments, de telle façon que si un point est sur le $K^{\text{ème}}$ segment, son foyer est sur le $(K+1)^{\text{ème}}$, son second foyer sur le $(K+2)^{\text{ème}}$, etc.

Reprenons la figure 3, soit A un point quelconque de C, AM_1BN_1E une géodésique infiniment voisine de C, de sorte que B et E soient le 1$^{\text{er}}$ et le 2$^{\text{d}}$ foyers de A. Si $n > 1$, cette courbe est disposée comme sur la figure, et la seule condition à remplir pour que notre théorème soit applicable, c'est que l'angle sous lequel cette courbe coupe C en A soit égal à l'angle sous lequel elle coupe C en E.

Soit u la valeur de u au point A et soit $F(u)$ le rapport de l'angle sous lequel les deux courbes se coupent en E à l'angle sous lequel elles se coupent en A. Le rapport $F(u)$ est toujours positif, et la question est de savoir si l'on peut disposer de u de telle façon que

$$F(u) = 1.$$

Soit $u = u_0$; l'équation de la géodésique AM_1BN_1E se réduit alors à

$$v = e^{\alpha u}\phi(u)$$

(ou à $v = e^{\alpha' u}\phi_1(u)$ dans le cas des géodésiques instables de la 3$^{\text{ème}}$ catégorie). Elle vient couper C successivement en $A_0, A_1, \cdots$, points correspondant aux arguments $u_0, u_1, \cdots$; de sorte que A_{k+1} est le foyer de A_k, et l'intersection a lieu sous les angles

$$\theta(u_0), \theta(u_1) \cdots,$$

en posant

$$\theta(u) = e^{\alpha u}[\alpha\phi(u) + \phi'(u)].$$

On aura donc

$$F(u_0) = \frac{\theta(u_2)}{\theta(u_0)}, \quad F(u_2) = \frac{\theta(u_1)}{\theta(u_2)}, \quad \cdots, \quad F(u_{2n-2}) = \frac{\theta(u_{2n})}{\theta(u_{2n-2})};$$

d'où

$$F(u_0)F(u_2)\cdots F(u_{2n-2})=\frac{\theta(u_{2n})}{\theta(u_0)}=\frac{\theta(u_0+2U)}{\theta(u_0)}=e^{2aU}>1,$$

ce qui prouve que l'un au moins des facteurs

$$F(u_0),\ F(u_2),\ \cdots,\ F(u_{2n-2})$$

est plus grand que 1, et par conséquent que $F(u)$ peut devenir >1. Soit maintenant $u=u_0'$, la géodésique se réduit à

$$v=e^{-au}\psi(u)$$

et l'on a

$$F(u_0')=\frac{\omega(u_2')}{\omega(u_2')},\quad F(u_2')=\frac{\omega(u_4')}{\omega(u_2')},\quad \cdots,\quad F(u_{2n-2}')=\frac{\omega(u_{2n}')}{\omega(u_{2n-2}')},$$

en posant

$$\omega(u)=e^{-au}[\psi'(u)-\alpha\psi(u)];$$

d'où,

$$F(u_0')F(u_2')\cdots F(u_{2n-2}')=\frac{\omega(u_{2n})}{\omega(u_0)}=\frac{\omega(u_0+2U)}{\omega(u_0)}=e^{-2aU}<1,$$

ce qui prouve que l'un au moins des facteurs

$$F(u_0'),\ F(u_2'),\ \cdots,\ F(u_{2n-2}')$$

est <1, et par conséquent que $F(u)$ peut devenir <1.

Le rapport $F(u)$ étant une fonction continue et périodique de u et pouvant devenir <1 et >1, pourra devenir égal à 1.

Ce qui montre encore que C n'est pas un vrai minimum.

Ainsi pour que la géodésique fermée C, soit la plus courte de toutes les courbes fermées sans point double qui enveloppent une aire de courbure totale 2π, il faut, si elle est stable que $\alpha U/i$ soit plus petit que 2π et si elle est instable que le nombre $2n$ des foyers limites soit égal à 2.

Ces conditions nécessaires sont-elles suffisantes? Je n'ai pas besoin de traiter ici la question que les procédés ordinaires du calcul des variations permettraient de résoudre.

THE CLASSIFICATION OF QUADRICS*

BY

T. J. I'A. BROMWICH

In a former number of the Transactions† Mr. J. L. COOLIDGE has given a classified list of the possible species of quadrics in hyperbolic space; his method is based on that used by CLEBSCH for discussing the mutual relations of two quadrics. It seems desirable, however, that the classification should be associated with the algebraic problem of reducing two quaternary quadratic forms to canonical types; the solution of this problem is, of course, based on WEIERSTRASS's theory of *invariant factors*.‡ One obvious advantage of using invariant factors as a basis of classification is that they can be determined *without any preliminary reduction* of the equation to the quadric; indeed, it is generally necessary to find the factors before we can begin the reduction. Another advantage may be found in the fact that the canonical types given by WEIERSTRASS are the most compact possible.§

The first classification of two quadrics by means of invariant factors is due to SYLVESTER.‖ From the geometrical point of view, an elaborate discussion was carried out, independently of WEIERSTRASS's methods, by PAINVIN;¶ PAINVIN touches on the question of *reality*, but only in the simplest case (when the invariant factors are *linear*). According to MUTH, the application to quadrics

* Presented to the Society at the St. Louis meeting, September 17, 1904. Received for publication March 4, 1904.

† Transactions of the American Mathematical Society, vol. 4 (1903), p. 161.

‡ *Elementartheiler;* see his paper on bilinear and quadratic forms, Berliner Monatsberichte, 1868, p. 310 = *Gesammelte Werke*, vol. 2, p. 19. For a bibliography, consult a review of MUTH'S *Elementartheiler* (Leipzig, 1899), in the Bulletin of the American Mathematical Society, vol. 7 (1901), p. 308. An elementary account of the application of this theory to the reduction of two quadratic forms will be found in a paper by the writer in the Quarterly Journal of Mathematics, vol. 33 (1901), p. 85.

§ As we shall see later, some of COOLIDGE'S types contain superfluous coefficients, which may be removed by suitable linear substitutions.

‖ Philosophical Magazine, ser. 4, vol. 1 (1851), pp. 119, 295; this paper is, of course, earlier than WEIERSTRASS'S. But it is not complete, as no method for determining the reducing substitutions is given, and thus it is not certain that the classification is exhaustive. In fact the two "singular cases" are omitted; i. e., when the quadrics are cones having (i) a common tangent-plane, which they touch along a common generator (ii) a common vertex.

¶ Nouvelles Annales de Mathématiques, ser. 2, vol. 7 (1868), pp. 481, 529; vol. 8 (1869), pp. 49, 145, 193. The last section of his paper contains a list of the chief results obtained.

of WEIERSTRASS's results (and of KRONECKER's,* on the "singular case") has been fully carried out by KILLING.†

WEIERSTRASS's method of determining the reducing substitutions does not show how to modify them in case we restrict ourselves to real quadratic forms and real linear substitutions. This difficulty was first surmounted by KLEIN in his celebrated memoir on the classification of line-complexes of the second degree‡; KLEIN shows there that if one of the quadratic forms has a known type, and if we are restricted to real forms (and real substitutions), then the possible invariant-factors are restricted by a certain relation.‖ KLEIN gives full details of all possible cases, if the number of variables is six, and the signature of the one known quadratic form is zero (see §1 below, for definition of *signature*).

BÔCHER§ has carried out a similar classification for the case of five variables, with a quadratic form of signature three.

In what follows we shall consider the case of four variables, the given signature being two (§2), four (§3), or three (§4).

§1. *Preliminary results on* REAL *transformations of quadratic forms.*

Any real quadratic form can be expressed in the shape

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2$$

(where $x_1, x_2, \cdots, x_n$ are real and linearly independent) in an infinity of ways; but, however the reduction is effected, the number of positive coefficients and the number of negative coefficients (in $a_1, a_2, \cdots, a_n$) are both *fixed*. This is SYLVESTER's *law of inertia* (*Trägheitsgesetz*) of quadratic forms.¶

* KRONECKER published a number of papers on this case; a list would be too long for enumeration here. The earlier papers will be found in the second half of volume 1 of his works; the last paper (in which the question was finally settled) appeared in the Berliner Sitzungsberichte of 1890, 1891. Alternative methods have been given by the present writer, Proceedings of the London Mathematical Society, vol. 32 (1900), pp. 87, 117, 326.

† *Der Flächenbüschel zweiter Ordnung, Inauguraldissertation*, Berlin, 1872; unfortunately this memoir has not been accessible. It does not appear from MUTH's references, whether or no KILLING concerns himself with the question of reality.

‡ *Inauguraldissertation*, Bonn, 1868; reprinted in vol. 23 of the Mathematische Annalen, p. 539.

‖ Mathematische Annalen, vol. 23, pp. 561, 562; the theorem has been extended to "HERMITE's forms" by A. LOEWY, Crelle's Journal für die Mathematik, vol. 122 (1900), p 69; see also a paper by the writer, Proceedings of the London Mathematical Society, vol. 32 (1901), p. 349.

§ *Reihenentwickelungen der Potentialtheorie*, Leipzig, 1894; the classification (of cyclides) will be found in Kapitel 3. It is perhaps of interest to note that the results obtained in §2 below can also be used to classify the cyclic curves (i. e. bicircular quartics), mentioned by BÔCHER.

¶ Philosophical Magazine, ser. 4, vol. 2 (1852), p. 138; Philosophical Transactions of London, 1853, p. 407. BORCHARDT (Crelle's Journal für die Mathematik,

For our present purpose, the difference between these two numbers is the most useful constant; this difference we call the *signature* of the quadratic form, after FROBENIUS.*

We are now in a position to state KLEIN's theorem (given in the dissertation already cited): —

If A is a real quadratic form, whose signature is known to be m, then, whatever quadratic form B may be (provided that it is real), the characteristic determinant $|\lambda A - B|$ has at least m REAL *invariant factors of* ODD *degree.*†

In the present application, A is the absolute quadric (of hyperbolic space) and takes the form

$$x_1^2 + x_2^2 + x_3^2 - x_4^2$$

so that its signature is 2(i. e. $3 - 1$). Accordingly, whatever quadric B may be, the determinant $|\lambda A - B|$ has at least two real invariant-factors of odd degree, Now the degree of this determinant is 4, so that the two real invariant-factors must have degrees $1, 1$ or $3, 1$. Hence the only possible combinations of indices for the invariant-factors are given by the sets

$$1111;\ 211;\ 31.$$

It is now clear that complex roots can occur only in the first set; further, in this set there may be one pair of such roots (not more). Thus the four fundamental types are

I	II	III	IV
1111	11 11	211	31

vol. 53 (1857), p. 275) points out that JACOBI was acquainted with the result as early as 1847; but apparently JACOBI had no occasion to publish it.

The restrictions that $x_1, x_2, \cdots, x_n$ must be real and linearly independent are essential; for example, we have

$$(x+iy)^2 + (x-iy)^2 = 2x^2 - 2y^2, \text{ and } x^2 + y^2 + (x+y)^2 + (x-y)^2 = 3x^2 + 3y^2.$$

*Crelle's Journal für die Mathematik, vol. 114 (1894), p. 187: the paper was also published in the Berliner Sitzungsberichte of that year.

†This theorem requires modification in case $|A| = 0$. For instance, if there are three variables x, y, z and if $A = x^2 + y^2$, $B = 2yz$, the signature of A is 2, but there is only *one* invariant-factor of $|\lambda A + \mu B|$ which has an odd degree; the invariant-factors are actually λ, μ^2. Of course in KLEIN's work the case $|A| = 0$ is excluded owing to the nature of the problem.

The modified theorem may be stated as follows: *Let $|\lambda A + \mu B|$ have n_1 invariant-factors of the type $(\lambda + \mu\epsilon)^\gamma$, ϵ being real (including zero) and γ being odd; and let n_2 be the number of invariant-factors of the type μ^β where β is even; then*

$$n_1 + n_2 \geqq m.$$

In the illustration above we have $n_1 = 1$, $n_2 = 1$, and the theorem is verified. For a discussion of the last theorem, see A. LOEWY, Crelle's Journal für die Mathematik, vol. 122 (1900), p. 61; BROMWICH, Proceedings of the London Mathematical Society, vol. 32 (1900), p. 162.

where, in case II, we have one pair of conjugate complex roots, indicated by the bar over the third and fourth indices.

We can now write down the standard types of terms which occur in the reduced* equations to the quadrics A, B; lists of these types will be found in the papers already cited and in MUTH's *Elementartheiler*.

Let c be the typical real root of the determinant $|\lambda A - B|$, then the standard types are given by the following table: —

Invariant-factor	$\lambda - c$	$(\lambda - c)^2$	$(\lambda - c)^3$
Terms in A	$\pm x^2$	$2xy$	$2xy + z^2$
Terms in B	$\pm cx^2$	$2cxy \pm x^2$	$c(2xy + z^2) + 2zx$

where x, y, z are all real and do not appear in any other terms.

Some modifications must, however, be introduced, in order to retain the regular type for the absolute A: and so we put (in the second and third cases) $x = \xi - \eta$, $2y = \xi + \eta$, obtaining the table: —

Invariant-factor	$(\lambda - c)^2$	$(\lambda - c)^3$
Terms in A	$\xi^2 - \eta^2$	$\xi^2 - \eta^2 + z^2$
Terms in B	$c(\xi^2 - \eta^2) \pm (\xi - \eta)^2$	$c(\xi^2 - \eta^2 + z^2) + 2z(\xi - \eta)$

Another change is necessary, when we have a pair of *complex* roots; if these are $\alpha + i\beta$, $\alpha - i\beta$, the corresponding reduced terms will be given by: —

Invariant-factor	$(\lambda - \alpha)^2 + \beta^2$
Terms in A	$\frac{1}{2}(x + iy)^2 + \frac{1}{2}(x - iy)^2 = x^2 - y^2$
Terms in B	$\frac{1}{2}[(\alpha - i\beta)(x + iy)^2 + (\alpha + i\beta)(x - iy)^2] = \alpha(x^2 - y^2) + 2\beta xy$

where x and y are real.

We have now sufficient material to write down all the types of quadrics by inspection.†

§ 2. *Canonical forms for the possible quadrics in hyperbolic space.*

Type I. Indices 1111.

$$A = y_1^2 + y_2^2 + y_3^2 - y_4^2,$$

$$B = c_1 y_1^2 + c_2 y_2^2 + c_3 y_3^2 - c_4 y_4^2, \qquad |B| = -c_1 c_2 c_3 c_4$$

where c_1, c_2, c_3, c_4 are roots of the determinantal equation in λ, $|\lambda A - B| = 0$.

* It is to be borne in mind that the invariant-factors can be calculated from the equation to B, given in any form; indeed they must, in general, be found before the reduction can be carried out.

† Since $|A| = -1$, there is here no reason to put down the types corresponding to the *identical* vanishing of $|\lambda A - B|$; nor to an infinite root of $|\lambda A - B| = 0$.

I(a). c_1, c_2, c_3, c_4 are all different, represented by the symbol [1111]. This case gives the *ellipsoid* and four *hyperboloids* found by COOLIDGE* (pp. 162, 163, l. c.).

I(b). $c_1 = c_2$ *or* $c_3 = c_4$, represented by [(11)11]. The case $c_1 = c_2$ gives the *spheroids* and *hyperboloids of revolution* (C., p. 168); $c_3 = c_4$ gives the *surfaces of translation* (C., § 4, p. 169).

I(c). $c_1 = c_2 = c_3$ *or* $c_1 = c_2 = c_4$; represented by [(111)1]. These cases appear as a *sphere* and an *equidistant surface*, respectively (C., p. 170).

I(d). $c_1 = c_2$ *and* $c_3 = c_4$; represented by [(11)(11)]. This gives the *analogue of Clifford's surface* in elliptic space (C., end of § 3, p. 169).

I(e). $c_1 = c_2 = c_3 = c_4$; represented by [(1111)]. This case is trivial, geometrically, as the quadric coincides with the absolute; it is, of course, algebraically distinct.

Type II. Indices 11 11.

$$A = y_1^2 + y_2^2 + y_3^2 - y_4^2,$$
$$B = c_1 y_1^2 + c_2 y_2^2 + \alpha(y_3^2 - y_4^2) + 2\beta y_3 y_4, \qquad |B| = -c_1 c_2 (\alpha^2 + \beta^2),$$

where the roots of $|\lambda A - B| = 0$ are c_1, c_2, $\alpha + i\beta$, $\alpha - i\beta$.

II(a). c_1, c_2 are unequal; represented by [11 11]. This is the *surface with imaginary centers* given by COOLIDGE (C., p. 165). But it may be well to point out that the apparently more general expression given by him is not actually more general than this (it contains five coefficients, one of which can be removed by linear transformations)†; and further, his equation has the disadvantage that its coefficients cannot be put down directly from the determinant $|\lambda A - B|$, before reduction.

II(b). $c_1 = c_2$; represented by [(11)11]. This is a *corresponding surface of revolution* (C. § 3, sub-case 4); here COOLIDGE's equation has the right number of arbitrary coefficients, but again they are not those naturally suggested by the determinant $|\lambda A - B|$; namely, $\alpha = \frac{1}{2}(a_3 - 1)$, $\beta = \frac{1}{2}[b^2 - (a_3 - 1)^2]^{\frac{1}{2}}$.

* I do not stay to repeat COOLIDGE'S sub-classification, which depends on the signs and relative magnitudes of the c's; and, in future, I quote his paper by the letter C.

† More exactly, two of them can be made equal, thus: Let p be that root of the quadratic $p + p^{-1} = 4b(a_3 + a_4)^{-1}$, which is numerically less than 1. Write

$$\alpha = \tfrac{1}{2}(a_3 - a_4), \qquad \beta = \tfrac{1}{4}(a_3 + a_4)(1 - p^2)p^{-1},$$
$$y_3 = (x_3 - px_4)(1 - p^2)^{-\frac{1}{2}}, \qquad y_4 = (px_3 - x_4)(1 - p^2)^{-\frac{1}{2}};$$

then we find

$$x_3^2 - x_4^2 = y_3^2 - y_4^2, \qquad a_3 x_3^2 - 2bx_3 x_4 + a_4 x_4^2 = \alpha(y_3^2 - y_4^2) + 2\beta y_3 y_4;$$

where

$$\alpha = \tfrac{1}{2}(a_3 - a_4) \text{ and } \beta = [b^2 - \tfrac{1}{4}(a_3 + a_4)^2]^{\frac{1}{2}}.$$

Of course p is real, because of the condition $4b^2 > (a_3 + a_4)^2$.

It might be thought that we could have cases to correspond to I(c), I(d); but this is not so, on the ground of our restriction to reality.

Type III. Indices 211.

$$A = y_1^2 + y_2^2 + y_3^2 - y_4^2,$$
$$B = c_1 y_1^2 + c_2 y_2^2 + c_3(y_3^2 - y_4^2) \pm (y_3 - y_4)^2, \qquad |B| = -c_1 c_2 c_3^2,$$

the roots of $|\lambda A - B| = 0$ being $\lambda = c_1, c_2, c_3, c_3$.

III(a). c_1, c_2, c_3 all different; represented by [211]. This gives the *elliptic and hyperbolic paraboloids** (C., p. 166).

III(b). $c_2 = c_3$; represented by [(21)1]. This is the *horocyclic paraboloid.* (C., p. 167).

III(c). $c_1 = c_2 = c_3$; represented by [(211)]. This is the *horocyclic surface of zero curvature* (C., p. 170); in this case the arbitrary sign must be positive, to avoid an ideal surface.

III(d). $c_1 = c_2$; represented by [2(11)]. This gives the *paraboloids of revolution* (C., p. 169).

Type IV. Indices 31.

$$A = y_1^2 + y_2^2 + y_3^2 - y_4^2,$$
$$B = c_1 y_1^2 + c_2(y_2^2 + y_3^2 - y_4^2) + 2y_2(y_3 - y_4), \qquad |B| = -c_1 c_2^3,$$

the roots of $|\lambda A - B| = 0$, being c_1, c_2, c_2, c_2.

IV(a). $c_1 \gtrless c_2$; represented by [31]. This is a surface to which COOLIDGE has not assigned a name; it is the first on p. 167 (C.). The equation given there contains two superfluous constants a_2, a_3; a_2 may be replaced by 1, and a_3 by 0, without loss of generality.†

* Again I do not stop to give the sub-classifications. But it may be well to point out that the coefficients a_3, a_4 in COOLIDGE's equation are not invariants; the essentials are $c_3 = \frac{1}{2}(a_3 + a_4)$ and the *sign* of $(a_3 - a_4)$. For if $a_3 - a_4 = 2eq^2$, where $e = \pm 1$, and q is real, we can write $y_3 - y_4 = q(x_3 - x_4)$, $y_3 + y_4 = (x_3 + x_4)/q$ and then $y_3^2 - y_4^2 = x_3^2 - x_4^2$, while

$$a_3 x_3^2 + (a_4 - a_3) x_3 x_4 - a_4 x_4^2 = c_3(y_3^2 - y_4^2) + e(y_3 - y_4)^2.$$

Similarly in COOLIDGE's type for III(b), the constant a is superfluous, except as to sign; in the type for III(c), a must be positive and may be put equal to $+1$.

† To prove this statement, call COOLIDGE's equation B; A being the equation to the absolute. Then

$$B - A = (a_1 - 1)x_1^2 + (x_3 - x_4)[2a_2 x_2 + a_3(x_3 - x_4)] = (a_1 - 1)x_1^2 + 2y_2(y_3 - y_4),$$

provided that $2a_2 y_2 = 2a_2 x_2 + a_3(x_3 - x_4)$, $y_3 - y_4 = a_2(x_3 - x_4)$. Now substitute in A for x_2 in terms of y_2, y_3, y_4; thus we find

$$A = x_1^2 + y_2^2 + (y_3 - y_4)\left[-\frac{a_3 y_2}{a_2^2} + \frac{a_3^2}{4a_2^4}(y_3 - y_4) + \frac{1}{a_2}(x_3 + x_4)\right].$$

Call the last expression in square brackets $(y_3 + y_4)$, then

$$A = x_1^2 + y_2^2 + y_3^2 - y_4^2, \qquad B - A = (a_1 - 1)x_1^2 + 2y_2(y_3 - y_4).$$

IV(b). $c_1 = c_2$; represented by [(31)]. This is the third surface considered on p. 167 (C.); in the equation there given we may take $r = 2$ without loss of generality.

We have now exhausted all possible quadric surfaces, in hyperbolic space. There are thus *twelve* fundamental classes, represented by the symbols given in the table

	a	b	c	d
I	1111	(11)11	(111)1	(11)(11)
II	1111	(11)11		
III	211	(21)1	(211)	2(11)
IV	31	(31)		

This classification can be completely carried out by an examination of $|\lambda A - B|$, without any preliminary reduction. But, for the further subdivisions, a certain amount of reduction is necessary; as it is not possible to foretell in I which of the four roots c_1, c_2, c_3, c_4 is to be associated with the negative term in A. More can be done in III; for c_3 is necessarily associated with the negative square; but here the sign of $(y_3 - y_4)^2$ cannot be foretold. In cases II and IV we can obtain complete information from the determinant; as may be seen by considering the reduced forms.

§3. *Quadrics in elliptic space.*

The absolute is here a positive quadratic form and its signature is 4; hence the only indices possible are 1111, and the roots c_1, c_2, c_3, c_4 are all *real*.* Thus we are restricted to the possibilities given by the row numbered I in the table at the end of §2.

Hence every quadric in elliptic space can be reduced to the form

$$c_1 y_1^2 + c_2 y_2^2 + c_3 y_3^2 + c_4 y_4^2 = 0$$

the absolute being $y_1^2 + y_2^2 + y_3^2 + y_4^2 = 0$; the coefficients c_1, c_2, c_3, c_4 are all real, being the roots of $|\lambda A - B| = 0$.

The details of the cases are:

I(a). [1111]; *surfaces not of revolution*, ruled or not according as the product $c_1 c_2 c_3 c_4$ is positive or negative.

I(b). [(11)11]; *surfaces of revolution* ($c_1 = c_2$), ruled or not according as $c_3 c_4$ is positive or negative.

I(c). [(111)1]; *spheres* (not ruled), $c_1 = c_2 = c_3$.

*This fact follows also from two theorems due to WEIERSTRASS (Berliner Monatsberichte, 1858, p. 207; *Gesammelte Werke*, vol. 1, p. 233; the theorems are given in §§3, 4).

I(d). [(11)(11)]; CLIFFORD'S *surfaces*, (ruled) $c_1 = c_2$, $c_3 = c_4$.

It may be remarked that there are no hyperboloids in any of these cases.

§ 4. *Quadrics in parabolic (Euclidean) space.*

The classification of quadrics in Euclidean space is carried out in most of the ordinary text-books on solid geometry, and details need hardly be given here; but a few remarks may be added, to indicate the contrast to the cases considered in §§ 2, 3.

In parabolic space the absolute is degenerate; its equation can be expressed only in terms of plane-coördinates u_1, u_2, u_3, u_4. If then $A = 0$ is the absolute, it will be found that the determinant $|A| = 0$, and that the signature of A is 3. Thus, if $B = 0$ is any quadric (in plane-coördinates) the determinant* $|\lambda A + \mu B|$ has

either (i) four linear invariant-factors $\lambda + \mu c_1$, $\lambda + \mu c_2$, $\lambda + \mu c_3$, μ;

or (ii) two linear invariant-factors $\lambda + \mu c_1$, $\lambda + \mu c_2$ and one μ^2;

where, in each case, c_1, c_2, c_3 (or c_1, c_2) are *real*.

From these facts we see that, by applying WEIERSTRASS'S methods (in conjunction with a slight modification† so as to include the factors μ, or μ^2), the reduced types are

(i) $$\begin{cases} A = v_1^2 + v_2^2 + v_3^2, \\ B = c_1 v_1^2 + c_2 v_2^2 + c_3 v_3^2 + c_4 v_4^2, \end{cases}$$

(ii) $$\begin{cases} A = v_1^2 + v_2^2 + v_3^2, \\ B = c_1 v_1^2 + c_2 v_2^2 + 2f v_3 v_4. \end{cases}$$

Here the coefficients c_4, f cannot be found from the determinant $|\lambda A + \mu B|$; it is not usually necessary to find them merely for the purpose of classification.‡ But they must be found if we require the precise size of the surface B: one method has been given by the present writer elsewhere.§

Another possibility may occur if $|B| = 0$; it may happen that

$$|\lambda A + \mu B| = 0,$$

* Either by the modified form of KLEIN'S theorem given above; or by a theorem due to GUNDELFINGER which is given in his edition of HESSE'S lectures on geometry and also in his lectures on conics (p. 67). See MUTH'S *Elementartheiler*, p. 180.

† Proceedings of the London Mathematical Society, vol. 32 (1900), p. 158.

‡ To distinguish between ruled and non-ruled surfaces, the sign of $|B|$ is sufficient

§ Proceedings of the Cambridge Philosophical Society, vol. 10 (1901), p. 358 (see § 3); see also BRÜCKEL'S collection of GUNDELFINGER'S formulae in Crelle's Journal für die Mathematik, vol. 119 (1898), pp. 210, 313.

identically. If so, the reduced forms are

$$\text{(iii)} \qquad \begin{cases} A = v_1^2 + v_2^2 + v_3^2, \\ B = c_1 v_1^2 + c_2 v_2^2 + c_3 v_3^2, \end{cases}$$

B being a conic in the plane at infinity.

But if a quadric given in point-coördinates is degenerate the method just explained breaks down entirely; for the transformation of such an equation to plane-coördinates is no longer possible. * To meet this difficulty an alternative method of reduction is explained in §§ 1, 2 of the paper just quoted; and a simple general method of classification is obtained. But the process by which the result is obtained is not very easy; and it may therefore be worth while to state the result in order to make it accessible to those who do not care to go through the details which are unavoidable in finding the reducing substitution. Let $R = 0$ be the given quadric, expressed in point-coördinates x_1, x_2, x_3, x_4; $S = 0$ any sphere, the coefficients of S having been so chosen that the expression $S/(p_1x_1 + p_2x_2 + p_3x_3 + p_4x_4)^2$ represents the square of the tangent from (x_1, x_2, x_3, x_4) to the sphere, where $p_1x_1 + p_2x_2 + p_3x_3 + p_4x_4 = 0$ is the plane at infinity. Write Δ_0 for the determinant $|\lambda R - S|$: and Δ_1 for the symmetrical determinant formed by adding a fifth row and column $(p_1, p_2, p_3, p_4, 0)$ to Δ_0. Then the elements of classification are †:

(i) the factors of Δ_1; say $1 - \lambda c_1$, $1 - \lambda c_2$, $1 - \lambda c_3$; where c_1, c_2, c_3 are all real and some may be zero.

(ii) the first term (say $k\lambda^\alpha$) in the expansion of $(-\Delta_0/\Delta_1)$ in *descending* powers of λ.

When the quadric is transformed to orthogonal Cartesian coördinates (say y_1, y_2, y_3), the terms of the second degree can be reduced to

$$c_1y_1^2 + c_2y_2^2 + c_3y_3^2$$

and, as to the remaining terms, three cases are possible, according to the value of the integer α. To put the work in as short a form as possible, we assume (from the results of the elementary investigations) that R, S must take the forms

$$R = c_1y_1^2 + c_2y_2^2 + c_3y_3^2 + 2fy_3 + c_4,$$

$$S = y_1^2 + y_2^2 + y_3^2 + 2(l_1y_1 + l_2y_2 + l_3y_3) + m,$$

* More precisely, the transformed equation either vanishes identically or fails to give sufficient information. It should be observed that the corresponding difficulty does not arise in hyperbolic or elliptic space; for the absolute may be expressed either in point or in plane-coördinates.

† It is the second of these which constitutes the chief simplification introduced by this method of classification.

where, in R, *either* f is zero, *or* both c_3, c_4 are zero. It is then easy to see that

$$\Delta_1 = -M^2(\lambda c_1 - 1)(\lambda c_2 - 1)(\lambda c_3 - 1),$$

$$\Delta_0 = M^2[(\lambda c_1 - 1)(\lambda c_2 - 1)(\lambda c_3 - 1)(\lambda c_4 - m) - l_1^2(\lambda c_2 - 1)(\lambda c_3 - 1) - l_2^2(\lambda c_3 - 1)(\lambda c_1 - 1) - (f\lambda - l_3)^2(\lambda c_1 - 1)(\lambda c_2 - 1)],$$

M being the determinant of a certain linear substitution. From this we find that c_1, c_2, c_3 are the letters obtained in (i) from the factors of Δ_1; also that

$$-\frac{\Delta_0}{\Delta_1} = \lambda c_4 - m - \frac{l_1^2}{\lambda c_1 - 1} - \frac{l_2^2}{\lambda c_2 - 1} - \frac{(f\lambda - l_3)^2}{\lambda c_3 - 1}.$$

Now, according to (ii), $k\lambda^\alpha$ is the first term of this expression when expanded in powers of $1/\lambda$; thus, we have three cases:

(1) $f = 0$; then $k\lambda^\alpha = \lambda c_4$; so that $k = c_4$, $\alpha = 1$,

(2) $f = 0$, $c_4 = 0$; here $\alpha < 1$,

(3) $c_3 = 0$, $c_4 = 0$; then $k\lambda^\alpha = (f\lambda)^2$; so that $k = f^2$, $\alpha = 2$.

Thus, classifying by means of the index α, we may write the results

$$\alpha < 1, \qquad R = c_1y_1^2 + c_2y_2^2 + c_3y_3^2,$$

$$\alpha = 1, \qquad R = c_1y_1^2 + c_2y_2^2 + c_3y_3^2 + k,$$

$$\alpha = 2, \qquad R = c_1y_1^2 + c_2y_2^2 + 2k^{\frac{1}{2}}y_3,$$

giving a complete classification of all quadrics* in point-coördinates.

QUEEN'S COLLEGE, GALWAY,
February 22, 1904.

ADDITION. Mr. COOLIDGE has kindly recalled to my memory the fact that a classification of quadrics, in hyperbolic and elliptic space, had been previously given by P. BARBARIN.† In this classification of hyperbolic space, the classes I, II, III, are the same as the classes which bear these numbers in § 2 above; but class IV is left out. The omission is due to an oversight in connection with the transformations given at the top of p. 113 (l. c.), § 28; in fact if the absolute is $x^2 + y^2 + z^2 - u^2 = 0$, the quadric $z^2 + 2x(y - u) = 0$ has only one real principal plane ($z = 0$); and this quadric cannot be reduced to BARBARIN's typical equation (40), $S_1X^2 + S_2y^2 + S_2Z^2 - 2PUX = 0$. There is a similar difficulty with any quadric belonging to class IV of § 2.

* It may be worth while to point out that if the quadric is degenerate one or more of c_1, c_2, c_3 may be zero; but this involves no change in the calculation of c_4 and f.

† *Études de Géometrie Non-Euclidienne*, Bruxelles, 1900, pp. 101–126; this is an extract (published separately) from vol. 60 of the Mémoires couronnés et autres Mémoires publiés par l'Académie royale de Belgique.

It is to be noticed that BARBARIN's nomenclature differs in some respects from COOLIDGE's, which has been used above: the former gives the generic name *paraboloid* to class II, and *horiquadric* to class III. Although there is an advantage in giving a name to each class, yet it does not appear to me that the name paraboloid is specially suited to class II, the quadrics of which have imaginary centers.

In conclusion, the classification of conics in the hyperbolic plane, carried out on the lines of §§ 1, 2, is:

I(a): $[111]$; $c_1y_1^2 + c_2y_2^2 - c_3y_3^2 = 0$.

(b): $[(11)1]$; $c_1(y_1^2 + y_2^2) - c_3y_3^2 = 0$, and $c_1y_1^2 + c_2(y_2^2 - y_3^2) = 0$.

(c): $[(111)]$; $y_1^2 + y_2^2 - y_3^2 = 0$ (the *absolute*).

II: $[\bar{1}\bar{1}1]$; $c_1y_1^2 + \alpha(y_2^2 - y_3^2) + 2\beta y_2y_3 = 0$.

III(a): $[21]$; $c_1y_1^2 + c_2(y_2^2 - y_3^2) \pm (y_2 - y_3)^2 = 0$.

(b): $[(21)]$; $c_1(y_1^2 + y_2^2 - y_3^2) \pm (y_2 - y_3)^2 = 0$.

IV: $[3]$; $c_1(y_1^2 + y_2^2 - y_3^2) + 2y_1(y_2 - y_3) = 0$.

In each case the absolute is $y_1^2 + y_2^2 - y_3^2 = 0$; and, according to KLEIN's theorem, there is at least one real invariant factor of odd degree.

The numbering of classes I, II, III corresponds to that of BARBARIN (l. c. pp. 39–57), who omits class IV; in fact the method given on p. 33 for determining the rotation θ fails if applied to the conics IV; and these conics cannot be reduced to the form contained in BARBARIN's equation (20_1) on p. 40, which is the basis of his subsequent work.

It is of course possible to give a classification of conics in the elliptic and in the parabolic planes, by a suitable modification of the work above; but the necessary changes are so obvious that it is hardly worth while to enumerate them.

March 15, 1905

ON DIFFERENTIAL INVARIANTS*

BY

JOSEPH EDMUND WRIGHT

Introduction.

In the consideration of differential equations there arise expressions, such as for instance, the JACOBI-POISSON alternant of two first order partial equations, which are in their nature invariantive to all contact transformations. An important problem immediately presents itself, namely, the obtaining of all invariants of this type, that is of all invariants, with respect to contact transformations, of differential expressions.

In this paper are obtained all such invariants of a restricted type.

The restrictions are the following:—

(1) The only expressions considered are: (*a*) expressions of the first order with m dependent and n independent variables; (*b*) expressions of the second order with one dependent variable.

(2) The invariants are only of the first order, that is to say, they involve only first derivatives of the differential expressions.

The variables assumed to occur in case (*a*) are

$$x_1, \cdots, x_n \text{ (the independent variables)},$$

$$z_1, \cdots, z_m \text{ (the dependent variables)},$$

$$p_k^i \equiv \frac{\partial z_i}{\partial x_k} \qquad (i=1, \cdots, m;\ k=1, \cdots, n),$$

$$p_{kl}^i \equiv \frac{\partial^2 z_i}{\partial x_k \partial x_l} \qquad (i=1, \cdots, m;\ k, l=1, \cdots, n),$$

and

$$\frac{\partial f_\lambda}{\partial x_k}, \quad \frac{\partial f_\lambda}{\partial z_i}, \quad \frac{\partial f_\lambda}{\partial p_k^i} \qquad (\lambda=1, \cdots, r;\ i=1, \cdots, m;\ k=1, \cdots, n),$$

where $f_\lambda(x, z, p_k^i)$, $(\lambda=1, \cdots, r)$, are the differential expressions considered.

In addition the variables dx_k, dz_i, dp_k^i will be assumed to enter, subject to the conditions

$$dz_i = \sum_{k=1}^{n} p_k^i dx_k \qquad (i=1 \cdots m),$$

$$dp_k^i = \sum_{l=1}^{n} p_{kl}^i dx_l \qquad (i=1, \cdots, m;\ k=1, \cdots, n).$$

* Presented to the Society April 29, 1905. Received for publication November 12, 1904.

In invariants corresponding to case (b) the variables are the same, with the exception that m is 1, and that there are additional variables

$$\frac{\partial f_\lambda}{\partial p_{kl}} \qquad (\lambda = 1, \cdots, r; k, l = 1, \cdots, n),$$

where $f_\lambda(x, z, p_k, p_{kl})$ are the differential expressions considered.

The method used is a modification of that given by LIE in his paper "Ueber Differentialinvarianten," * in which invariants for certain simpler types of infinitesimal transformations are obtained. In this paper LIE shows (p. 566) that by the method there outlined, a series of invariants may be obtained satisfying a system of linear differential equations which form a complete system.

In the paper mentioned LIE suggests a problem connected with invariants of surfaces, the class desired being that which does not change owing to "deformation" of the surface.

ŻORAWSKI † attempted the solution of this question and found a class of such invariants.

FORSYTH, ‡ in 1903 attacked the more general question of invariants due to a purely arbitrary point transformation performed on the surface, and also obtained in this manner the differential invariants of space. In his paper certain modifications are made on the LIE-ŻORAWSKI method, one modification being that he sought for relative, as well as absolute, invariants. The method as modified by FORSYTH will be used here.

The invariant sought will, therefore, be such that if F denote its expression in the original, F_1 in the transformed variables, we shall have

$$F_1 = \Omega F,$$

where Ω is a function depending only on the transformation.

Now if the transformation were a general one in the variables considered, it is well known that Ω would be some power of the Jacobian of the transformation. But the transformation is not perfectly general.

In fact, in our most general case, the Jacobian of X_k, Z_i, P^i_k, P^i_{kl} with reference to x_k, z_i, p^i_k, p^i_{kl}, where capitals denote transformed variables, and where $\{k, l = 1, 2, \cdots n; i = 1, 2 \cdots m\}$ breaks up into two factors, the first of which is

$$J_1 \equiv J\left(\frac{X_1 X_2 \cdots X_n Z_1 \cdots Z_m P'_1 P'_2 \cdots P^m_n}{x_1 x_2 \cdots p^m_n}\right),$$

and the second

$$J_2 \equiv J\cdot\left(\frac{\cdots P^i_{kl} \cdots}{\cdots p^i_{kl} \cdots}\right).$$

* LIE, Mathematische Annalen, vol. 24 (1884), pp. 537–578.

† ŻORAWSKI, Acta Mathematica, vol. 16 (1892–3), pp. 1–64.

‡ FORSYTH, Philosophical Transactions, ser. A., vol. 201 (1903), pp. 329–402, and ser. A., vol. 202 (1904), pp. 277–333.

Further, if the transformation is a point transformation, $\bar{J}_1$ breaks up into two factors,

$$J_0 \equiv J\left(\frac{X_1 X_2 \cdots X_n Z_1 \cdots Z_m}{x_1 x_2 \cdots x_n z_1 \cdots z_m}\right),$$

$$J_1 \equiv J\left(\frac{P'_1 P'_2 \cdots P'^m_n}{p'_1 p'_2 \cdots p'^m_n}\right).$$

Now if the number of dependent variables is greater than one, it may easily be shown that the most general contact transformation possible is an extended point transformation.

The discussion will be limited to those cases in which the factor Ω is of the form

$$J_0^{\mu_0} J_1^{\mu_1} J_2^{\mu_2} \cdots$$

when the number of dependent variables is greater than one, and

$$\bar{J}_1^{\mu_1} J_2^{\mu_2} \cdots$$

when there is only one dependent variable.

Now let F be an invariant of the type considered, and let an infinitesimal contact transformation be performed on F.

The condition for invariance is that, t being the parameter of the transformation,

$$F_1 = \Omega F, \qquad \text{or} \qquad \frac{dF}{dt} = F\frac{d\Omega}{dt}$$

when $d\phi/dt$ denotes the increment of ϕ due to the infinitesimal transformation and F_1 is F in the tranformed variables. Expressing the fact that this equation holds for all such transformations as considered, we obtain a complete system of linear differential equations, the solutions of which are the invariants desired.

In the course of the work the values of certain increments are required, and they will be given now, before we consider the various cases in detail.

The following notation is used throughout:—

$$\frac{dx_k}{dt} \equiv \xi_k, \qquad \frac{dz_i}{dt} \equiv \zeta_i, \qquad \frac{dp^i_k}{dt} \equiv \pi^i_k, \qquad \frac{dp^i_{kl}}{dt} \equiv \pi^i_{kl}, \qquad \theta_i \equiv \zeta_i - \sum_{k=1}^{n} p^i_k \xi_k,$$

$$\frac{\partial f}{\partial x_k} \equiv X_k, \qquad \frac{\partial f}{\partial z_i} \equiv Z_i,$$

$$\frac{\partial f}{\partial p^i_k} \equiv P^i_k, \qquad \frac{\partial f}{\partial p^i_{kl}} \equiv P^i_{kl},$$

where f is used to denote one of the forms considered. If it is desired to specify any one of the f's particularly, the notation

$$f_\lambda, \qquad X_{\lambda,k}, \qquad P^i_{\lambda,k}, \qquad P^i_{\lambda,kl},$$

etc., is used.

$\mathbf{S}$ denotes a summation taken over all the expressions f.

In addition,

$$dx_k \equiv a_k, \qquad dz_i \equiv b_i, \qquad dp_k^i \equiv c_k^i, \qquad dp_{kl}^i \equiv c_{kl}^i,$$

$$\frac{d}{dx_k} \equiv \frac{\partial}{\partial x_k} + \sum_{\lambda=1}^{m} p_k^\lambda \frac{\partial}{\partial z_\lambda} + \sum_{\lambda, l} p_{kl}^\lambda \frac{\partial}{\partial p_l^\lambda}.$$

Using this notation we have the following increments:

$$\frac{dJ_0}{dt} \equiv \sum_{i=1}^{m} \frac{\partial \theta_i}{\partial z_i} + \sum_{k=1}^{n} \frac{d\xi_k}{dx_k},$$

$$\frac{dJ_1}{dt} \equiv n \sum_{i=1}^{m} \frac{\partial \theta_i}{\partial z_i} - m \sum_{k=1}^{n} \frac{d\xi_k}{dx_k},$$

$$\frac{dJ_2}{dt} \equiv \tfrac{1}{2} n(n+1) \sum_{i=1}^{m} \frac{\partial \theta_i}{\partial z_i} - m(n+1) \sum_{k=1}^{n} \frac{d\xi_k}{dx_k}$$

when m is greater than unity, and

$$\frac{d\bar{J}_1}{dt} \equiv (n+1) \frac{\partial \theta}{\partial z},$$

$$\frac{dJ_2}{dt} \equiv (n+1) \sum_{\lambda=1}^{n} \frac{d}{dx_h} \left(\frac{\partial \theta}{\partial p_h} \right) + \tfrac{1}{2} n(n+1) \frac{\partial \theta}{\partial z}$$

when m is unity.

The quantities π_k^i, π_{hk}^i, etc., are readily obtained in terms of the ξ's and ζ's and their derivatives by the method given in LIE-ENGEL, *Theorie der Transformationsgruppen*, vol. 1, p. 544, *et seq.*

In the case, however, where m is unity the increment of the variables in an extended infinitesimal contact transformation may be expressed in a particularly simple manner.

The theorem is as follows:

Let there be an extended infinitesimal contact transformation in the n independent variables $x_1, x_2, \cdots, x_n$ and the dependent variable z, and let $p_{hkl}\cdots$ denote $\partial^r z / \partial x_h \partial x_k \partial x_l \cdots$, where there are r letters $h, k, l, \cdots$. Then the increment of $p_{hkl}\cdots$ due to the transformation is $(d^r \theta / dx_h dx_k dx_l \cdots)\delta t$, where θ, with the usual notation, is equal to

$$\zeta - \sum_{i=1}^{n} p_i \xi_i,$$

and $(d^r \theta / dx_h dx_k dx_l \cdots)$ denotes a total differentiation of θ in which the terms containing the highest derivatives of z are omitted.

For example, in the case when there are two independent variables,

$$\pi_1 = \frac{\partial \theta}{\partial x_1} + p_1 \frac{\partial \theta}{\partial z}, \qquad \pi_2 = \frac{\partial \theta}{\partial x_2} + p_2 \frac{\partial \theta}{\partial z},$$

$$\pi_{11} = \frac{\partial^2 \theta}{\partial x_1^2} + 2p_1 \frac{\partial^2 \theta}{\partial z\, \partial x_1} + p_1^2 \frac{\partial^2 \theta}{\partial z^2} + 2p_{11}\left(\frac{\partial^2 \theta}{\partial x_1 \partial p_1} + p_1 \frac{\partial^2 \theta}{\partial z \partial p_1}\right) + 2p_{12}\left(\frac{\partial^2 \theta}{\partial x_1 \partial p_2} + p_1 \frac{\partial^2 \theta}{\partial z \partial p_2}\right) + p_{11}^2 \frac{\partial^2 \theta}{\partial p_1^2} + 2p_{11}p_{12} \frac{\partial^2 \theta}{\partial p_1 \partial p_2} + p_{12}^2 \frac{\partial^2 \theta}{\partial p_2^2} + p_{11} \frac{\partial \theta}{\partial z}, \text{ etc.}$$

This theorem is known to be true * for the increments of the first derivatives of z, and it may be easily proved for higher derivatives by induction.

The increments of the quantities $X_{\lambda,\, k}$, $Z_{\lambda,\, i}$, $P^i_{\lambda,\, k}$, $P^i_{\lambda,\, hk}$, owing to the infinitesimal contact transformation are determined by the method given by FORSYTH in his paper already quoted.†

Using this method, we have the results

$$-\frac{dX_\lambda}{dt} = \sum_{k=1}^{n} \frac{\partial \xi_k}{\partial x_\lambda} X_k + \sum_{i=1}^{m} \frac{\partial \zeta_i}{\partial x_\lambda} Z_i + \sum_{ik} \frac{\partial \pi^i_k}{\partial x_\lambda} P^i_k + \sum_{ihk} \frac{\partial \pi^i_{hk}}{\partial x_\lambda} P^i_{hk},$$

$$-\frac{dZ_\lambda}{dt} = \sum_{k=1}^{n} \frac{\partial \xi_k}{\partial z_\lambda} X_k + \sum_{i=1}^{m} \frac{\partial \zeta_i}{\partial z_\lambda} Z_i + \sum_{ik} \frac{\partial \pi^i_k}{\partial z_\lambda} P^i_k + \sum_{ihk} \frac{\partial \pi^i_{hk}}{\partial z_\lambda} P^i_{hk},$$

$$-\frac{dP^\lambda_\mu}{dt} = \sum_{k=1}^{n} \frac{\partial \xi_k}{\partial p^\lambda_\mu} X_k + \sum_{i=1}^{m} \frac{\partial \zeta_i}{\partial p^\lambda_\mu} Z_i + \sum_{ik} \frac{\partial \pi^i_k}{\partial p^\lambda_\mu} P^i_k + \sum_{ihk} \frac{\partial \pi^i_{hk}}{\partial p^\lambda_\mu} P^i_{hk},$$

$$-\frac{dP^\lambda_{\mu\nu}}{dt} = \sum_{k=1}^{n} \frac{\partial \xi_k}{\partial p^\lambda_{\mu\nu}} X_k + \sum_{i=1}^{m} \frac{\partial \zeta_i}{\partial p^\lambda_{\mu\nu}} Z_i + \sum_{ik} \frac{\partial \pi^i_k}{\partial p^\lambda_{\mu\nu}} P^i_k + \sum_{ihk} \frac{\partial \pi^i_{hk}}{\partial p^\lambda_{\mu\nu}} P^i_{hk}.$$

The increments of the quantities a_k, b_i, c^i_k, etc., are readily calculated, for the transformation changes x_k into $x_k + \xi_k \delta t$, and therefore dx becomes $dx_k + d\xi_k \delta t$.

Hence

$$\frac{da_k}{dt} = \sum_{l=1}^{n} \frac{\partial \xi_k}{\partial x_l} a_l + \sum_{\lambda=1}^{m} \frac{\partial \xi_k}{\partial z_\lambda} b_\lambda + \sum_{l,\lambda} \frac{\partial \xi_k}{\partial p^\lambda_l} c^\lambda_l.$$

Similarly

$$\frac{db_i}{dt} = \sum_{l=1}^{n} \frac{\partial \zeta_i}{\partial x_l} a_l + \sum_{\lambda=1}^{m} \frac{\partial \zeta_i}{\partial z_\lambda} b_\lambda + \sum_{l,\lambda} \frac{\partial \zeta_i}{\partial p^\lambda_l} c^\lambda_l,$$

with similar expressions for the other increments of this type.

* See LIE-ENGEL, *Theorie der Transformationsgruppen*, vol. 2, p. 82, 252.

† FORSYTH, Philosophical Transactions, vol. 201, p. 337, 338.

§ 1.

We shall now proceed to the determination of invariants of expressions of the first order. The case in which there is only one dependent variable differs from the others in that the most general infinitesimal contact transformation is not an extended point transformation. This case will therefore be considered independently of the other.

Assuming F to be an invariant, and performing on it the infinitesimal transformation corresponding to a function

$$\theta \equiv \zeta - \sum_{k=1}^{n} p_k \xi_k,$$

we have the equation

$$\frac{dF}{dt} + \mu(n+1)\theta_z F = 0.$$

If the variables occurring in F are

$$x_1, \cdots, x_n, z, p_1, \cdots, p_n, X_{\lambda,1}, \cdots, X_{\lambda,n}, Z_\lambda, P_{\lambda,1}, \cdots, P_{\lambda,n} \quad (\lambda = 1 \cdots r),$$

this equation becomes on expansion,

$$\mu(n+1)\theta_z F + \left(\theta - \sum_{i=1}^{n} p_i \theta_{p_i}\right)\frac{\partial F}{\partial z} + \sum_{i=1}^{n}(\theta_{x_i} + p_i\theta_z)\frac{\partial F}{\partial p_i} - \sum_{i=1}^{n}\theta_{p_i}\frac{\partial F}{\partial x_i}$$

$$- \mathbf{S}\sum_{i=1}^{n}\frac{\partial F}{\partial X_i}\left\{\sum_{\rho=1}^{n} X_\rho(-\theta_{p_\rho x_i}) + \left(\theta_{x_i} - \sum_{\rho=1}^{n} p_\rho \theta_{p_\rho x_i}\right)Z + \sum_{\rho=1}^{n}(\theta_{x_\rho x_i} + p_\rho\theta_{x_i z})P_\rho\right\}$$

$$- \mathbf{S}\frac{\partial F}{\partial Z}\left\{\sum_{\rho=1}^{n} X_\rho(-\theta_{p_\rho z}) + \left(\theta_z - \sum_{\rho=1}^{n} p_\rho\theta_{p_\rho z}\right)Z + \sum_{\rho=1}^{n}(\theta_{x_\rho z} + p_\rho\theta_{zz})P_\rho\right\}$$

$$- \mathbf{S}\sum_{i=1}^{n}\frac{\partial F}{\partial P_i}\left\{\sum_{\rho=1}^{n} X_\rho(-\theta_{p_\rho p_i}) + \sum_{\rho=1}^{n}(-p_\rho\theta_{p_\rho p_i})Z\right.$$

$$\left. + \sum_{\rho=1}^{n}(\theta_{p_i x_\rho} + p_\rho\theta_{z p_i})P_\rho + \theta_z P_i\right\} = 0.$$

Now θ is a perfectly arbitrary function of the variables $x_1 \ldots x_n, z, p_1 \cdots p_n$ and the above equation must be satisfied for all values of θ.

Hence we may equate to zero the coefficients of the derivatives of θ, and thus obtain the system of linear partial equations which F must satisfy.

The system is the following:

From θ

$$F_z = 0. \tag{1}$$

From θ_{x_i},

$$F_{p_i} - \mathbf{S}ZF_{X_i} = 0 \qquad (i = 1 \cdots n). \tag{2}$$

From θ_{p_i},

$$F_{x_i} = 0 \qquad (i=1, \cdots, n). \tag{3}$$

From θ_z,

$$\mu(n+1)F + \sum_{i=1}^{n} p_i F_{p_i} - \mathbf{S} Z F_Z - \mathbf{S} \sum_{i=1}^{n} P_i F_{P_i} = 0. \tag{4}$$

From $\theta_{x_i x_k}$,

$$\mathbf{S} P_k F_{X_i} + \mathbf{S} P_i F_{X_k} = 0 \qquad (i=1, \cdots, n;\ k=1, \cdots, n). \tag{5}$$

From $\theta_{x_i z}$,

$$\mathbf{S} \sum_{\rho=1}^{n} p_\rho P_\rho F_{X_i} + \mathbf{S} P_i F_Z = 0 \qquad (i=1, \cdots, n). \tag{6}$$

From $\theta_{x_i p_k}$,

$$\mathbf{S}(X_k + p_k Z) F_{X_i} - \mathbf{S} P_i F_{P_k} = 0 \qquad (i=1, \cdots, n;\ k=1, \cdots, n). \tag{7}$$

From θ_{zz},

$$\mathbf{S}\left(\sum_{\rho=1}^{n} p_\rho P_\rho \right) F_Z = 0. \tag{8}$$

From θ_{zp_i},

$$\mathbf{S}(X_i + p_i Z) F_Z - \mathbf{S}\left(\sum_{\rho=1}^{n} p_\rho P_\rho \right) F_{P_i} = 0 \qquad (i=1, \cdots, n). \tag{9}$$

From $\theta_{p_i p_k}$,

$$\mathbf{S}(X_k + p_k Z) F_{P_i} + \mathbf{S}(X_i + p_i Z) F_{P_k} = 0. \qquad (i, k=1, \cdots, n). \tag{10}$$

It follows from these equations that F must be a function of the variables

$$A_{1,1}, \cdots, A_{1,n}, A_{2,1}, \cdots, A_{2,n}, \cdots, Z_1, Z_2, \cdots, Z_r, P_{1,1}, \cdots, P_{r,n},$$

where

$$A_{i,k} \equiv X_{i,k} + p_k Z_i \qquad (i=1, \cdots, r;\ k=1, \cdots, n).$$

If we modify the system of equations by assuming F to be a function of these variables only, it becomes

$$\mu(n+1)F = \mathbf{S} Z F_Z + \mathbf{S} \sum_{i=1}^{n} P_i F_{P_i}, \tag{11}$$

$$\mathbf{S} P_i F_{A_k} + \mathbf{S} P_k F_{A_i} = 0 \qquad (i, k=1, \cdots, n), \tag{12}$$

$$\mathbf{S} P_i F_Z = 0 \qquad (i=1, \cdots, n), \tag{13}$$

$$\mathbf{S} A_k F_{A_i} - \mathbf{S} P_i F_{P_k} = 0 \qquad (i, k=1, \cdots, n), \tag{14}$$

$$\mathbf{S} A_i F_Z = 0 \qquad (i=1, \cdots, n), \tag{15}$$

$$\mathbf{S} A_i F_{P_k} + \mathbf{S} A_k F_{P_i} = 0 \qquad (i, k=1, \cdots, n). \tag{16}$$

If we put aside for the present the first of these equations, the remaining equations form a complete system. The number of functionally independent solu-

tions is, however, not immediately deducible, as some of the equations may depend algebraically on the others. In particular, equations (13) and (15) show that if the number r of expressions f considered is not greater than $2n$, and if all the determinants of order r of the matrix

$$\begin{Vmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} & P_{1,1} & \cdots & P_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} & P_{2,1} & \cdots & P_{2,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ A_{r,1} & A_{r,2} & \cdots & A_{r,n} & P_{r,1} & \cdots & P_{r,n} \end{Vmatrix}$$

do not vanish, F_Z must be zero.

Further, there is no necessity to consider more than $2n+1$ expressions f, for if there were more than this number they could all be expressed by means of any $(2n+1)$ of them.

Suppose that there is only one expression f, then F_{A_i}, F_Z, F_{P_i} $(i = 1, \cdots, n)$ are all zero, and therefore it follows that there is no invariant of the type sought of a single expression f.

Next let there be two expressions f. Then F_{Z_1}, F_{Z_2} are zero unless

$$\frac{A_{1,1}}{A_{2,1}} = \frac{A_{1,2}}{A_{2,2}} = \cdots = \frac{P_{1,i}}{P_{2,i}} = \cdots .$$

But these relations are equivalent to the conditions that any equation $\phi(x_i z p_i) = 0$ which lies in involution with $f_1 = a$, shall lie in involution with $f_2 = b$, where a and b are arbitrary constants.

Now corresponding to a particular value of a there are ∞^{2n-1} characteristic strips which go to build up the integrals of the equation $f_1 = a$.

Hence taking account of all values of a, all the surface elements $(z x_i p_i)$ in space of $(n+1)$ dimensions are arranged in ∞^{2n} characteristic strips.

The condition given above is easily seen to be the same as the condition that $f_2 = b$ determines the same system of ∞^{2n} characteristic strips.

Exactly similarly in the case of r expressions f, there are obtained ∞^{2n} $(2n - r + 1)$-fold manifolds which are common to the first $(r-1)$ equations $f_1 = a_1, \cdots, f_{r-1} = a_{r-1}$, where $a_1, \cdots, a_{r-1}$ are arbitrary constants, and the conditions in virtue of which $F_{Z_1}, F_{Z_2}, \cdots, F_{Z_r}$ are not zero are the conditions that these manifolds should also satisfy $f_r = a_r$ for all values of the constant a_r.

We accordingly neglect this particular case, and then all the differential coefficients F_Z are zero, provided $r < 2n+1$.

The system of equations which F satisfies is readily integrated, and it is seen that F must be a function of the POISSON alternants

$$[f_\lambda f_\mu] \qquad (\lambda, \mu = 1, 2 \cdots r),$$

and in addition must satisfy equation (11). This equation merely expresses the

fact that F must be homogeneous of degree $\mu(n+1)$ in the quantities P, and F is therefore a homogeneous function of the quantities

$$[f_\lambda f_\mu] \qquad (\lambda, \mu = 1, 2 \cdots r).$$

There remains one case still to be considered, namely that in which we have $(2n+1)$ expressions f. In this case there exist invariants of the form $[f_\lambda f_\mu](\lambda, \mu = 1, 2, \cdots, 2n-1)$ and among these the only relations are those of the type $[f_\lambda f_\mu] + [f_\mu f_\lambda] = 0$.

We therefore have $n(2n+1)$ functionally independent solutions of our system of equations.

But returning to this system, we find that it consists of $2n^2+3n$ equations in addition to an equation which expresses a condition of homogeneity. There are $(2n+1)^2$ variables involved in the equations, and the equations are now algebraically independent. They therefore possess $(2n+1)^2-(2n^2+3n)$ functionally independent solutions. Of this number, $(2n^2+n+1)$, $n(2n+1)$ are accounted for, and therefore one solution still remains to be discovered.

It is readily seen that this solution is

$$\begin{vmatrix} Z_1 & A_{1,1} & A_{1,2}, \cdots & P_{1,1} & P_{1,2}, \cdots & P_{1,n} \\ Z_2 & A_{2,1} & \cdots & & & P_{2,n} \\ \vdots & & & & & \\ Z_{2n+1} & A_{2n+1,1} & & & & P_{2n+1,n} \end{vmatrix} \equiv J.$$

If we substitute in (11) we see that, if F is J, μ is equal to unity.

Collecting results we see that *the only functionally independent relative invariants of our type, of r expressions f, are the alternants* $[f_\lambda f_\mu]$ *if r is less than* $2n+1$, *and if r is equal to* $(2n+1)$ *there is one additional invariant, the Jacobian of the forms with respect to the variables involved in them.*

It is well known that the alternants $[f_\lambda f_\mu]$ are all invariants of the forms f.

The theorem that these are the only invariants of the type sought, has been given by LIE,* who, however, merely suggests the method of proof. Further, LIE has apparently overlooked the additional invariant which arises in connection with $(2n+1)$ forms though he must have been perfectly familiar with the fact that the Jacobian is an invariant.

§2.

Let us now consider expressions involving one dependent variable, n independent variables, and the derivatives of the dependent variable with respect to the independent ones of the first and second orders.

The variables involved are now

$$z,\ x_1,\ x_2,\ \cdots,\ x_n,\ p_1,\ p_2,\ \cdots,\ p_n,\ \cdots,\ p_{ik} \qquad (i,\ k = 1, 2 \cdots n).$$

* LIE, Mathematische Annalen, vol. 24 (1884), p. 578, and Göttingen Nachrichten (1872), pp. 478–479.

The invariant sought will be a function of these, and of the first derivatives of the expression with respect to the variables involved in them.

Assume that the infinitesimal contact transformation is determined as before by a function θ, and that F is the invariant. Then the equation satisfied by F is

$$\frac{dF}{dt} + \left\{ \mu_0(n+1)\frac{\partial\theta}{\partial z} + \mu_1\left[\frac{n(n+1)}{2}\frac{\partial\theta}{\partial z} + (n+1)\sum_{k=1}^{n}\frac{d}{dx_k}\left(\frac{\partial\theta}{\partial p_k}\right)\right]\right\} F = 0.$$

Expand this, and it becomes

$$F_z\frac{dz}{dt} + \sum_{r=1}^{n} F_{x_r}\frac{dx_r}{dt} + \sum_{r=1}^{n} F_{p_r}\frac{dp_r}{dt} + \sum_{\alpha\beta} F_{p_{\alpha\beta}}\frac{dp_{\alpha\beta}}{dt} + \mathbf{S}F_Z\frac{dZ}{dt} + \mathbf{S}\sum_{r=1}^{n} F_{X_r}\frac{dX_r}{dt}$$

$$+ \mathbf{S}\sum_{r=1}^{n} F_{P_r}\frac{dP_r}{dt} + \mathbf{S}\sum_{\alpha\beta} F_{P_{\alpha\beta}}\frac{dP_{\alpha\beta}}{dt} + \left[\lambda\theta_z + \mu\sum_{r=1}^{n}\frac{d}{dx_r}(\theta_{p_r})\right] F = 0,$$

where

$$\lambda = (n+1)\left(\mu_0 + \frac{n}{2}\mu_1\right),$$

$$\mu = (n+1)\mu_1.$$

As before, F is an invariant to all contact transformations, and therefore, if we substitute in the above equation the values of

$$\frac{dz}{dt}, \frac{dx_r}{dt}, \cdots, \frac{dZ}{dt}, \cdots$$

in terms of θ and its derivatives, and if we then equate to zero the coefficients of the various derivatives of θ, we obtain a system of linear differential equations which F must satisfy.

If we equate to zero the coefficients of

$$\theta,\ \theta_{x_1},\ \cdots,\ \theta_{x_n},\ \theta_{x_1x_1},\ \cdots,\ \theta_{x_nx_n},\ \theta_{p_1},\ \cdots,\ \theta_{p_n},$$

we obtain the following system of equations:

$$F_z = 0, \qquad F_{x_i} = 0 \qquad (i = 1, 2, \cdots, n),$$

$$F_{p_i} - \mathbf{S}ZF_{X_i} = 0 \qquad (i = 1, 2, \cdots, n),$$

$$F_{p_{ii}} - \mathbf{S}P_iF_{X_i} = 0 \qquad (i = 1, 2, \cdots, n),$$

$$F_{p_{ik}} - \mathbf{S}P_iF_{X_k} - \mathbf{S}P_kF_{X_i} = 0 \qquad (i, k = 1, 2, \cdots, n, i \neq k).$$

We therefore introduce the variables

$$A_i = X_i + p_iZ + \sum_{k=1}^{n} p_{ik}P_k \qquad (i = 1, 2, \cdots, n),$$

and then these equations show that F is a function of

$$A_1,\ \cdots,\ A_n,\ Z,\ P_1,\ \cdots,\ P_n,\ P_{11},\ P_{12},\ P_{nn},$$

only.

The increments of the A's due to the transformation are readily found, and we have

$$-\frac{dA_i}{dt} = -\sum_{k=1}^{n} A_k\left(\frac{d\theta_{p_k}}{dx_i}\right) + \sum_{jk}\left\{\frac{d}{dx_i}\left(\frac{d^2\theta}{dx_j dx_k}\right)\right\}P_{jk},$$

where $(d^2\theta/dx_\alpha dx_\beta)$ denotes as before the differential coefficient of θ with respect to x_α and x_β, in which z, p_r, $p_{\alpha\beta}$ are taken to be functions of the x's, the terms containing third derivatives of z being omitted, and d/dx_i is written for

$$\frac{\partial}{\partial x_i} + p_i\frac{\partial}{\partial z} + \sum_{k=1}^{n} p_{ik}\frac{\partial}{\partial p_k}.$$

It is easy to show that

$$\frac{d}{dx_i}\left(\frac{d^2\theta}{dx_\alpha dx_\beta}\right) = \frac{d}{dx_\alpha}\left(\frac{d^2\theta}{dx_i dx_\beta}\right) = \frac{d}{dx_\beta}\left(\frac{d^2\theta}{dx_i dx_\alpha}\right) = \left(\left(\frac{d^3\theta}{dx_i dx_\alpha dx_\beta}\right)\right),$$

where the last expression denotes $d^3\theta/dx_i dx_\alpha dx_\beta$ with the terms containing third and fourth derivatives of z omitted.

The increments expressed in the variables A, Z, P are therefore as follows:

$$-\frac{dA_i}{dt} = -\sum_{k=1}^{n} A_k\frac{d\theta_{p_k}}{dx_i} + \sum_{jk}\left(\left(\frac{d^3\theta}{dx_i dx_j dx_k}\right)\right)P_{jk},$$

$$-\frac{dZ}{dt} = -\sum_{k=1}^{n} A_k\frac{\partial^2\theta}{\partial p_k \partial z} + Z\theta_z + \sum_{k=1}^{n} P_k\frac{d\theta_z}{dx_k} + \sum_{jk}\frac{\partial}{\partial z}\left(\frac{d^2\theta}{dx_j dx_k}\right)P_{jk},$$

$$-\frac{dP_i}{dt} = -\sum_{k=1}^{n} A_k\frac{\partial^2\theta}{\partial p_i \partial p_k} + \sum_{k=1}^{n} P_k\frac{d\theta_{p_i}}{dx_k} + P_i\theta_z + \sum_{jk}\frac{\partial}{\partial p_i}\left(\frac{d^2\theta}{dx_j dx_k}\right)P_{jk},$$

$$-\frac{dP_{hk}}{dt} = \sum_{\alpha\beta}\frac{\partial}{\partial p_{hk}}\left(\frac{d^2\theta}{dx_\alpha dx_\beta}\right)P_{\alpha\beta}.$$

If F is taken as a function of $A_1, \cdots, A_n, Z, P_1, \cdots, P_n, P_{11}, P_{12}, \cdots, P_{nn}$, the equation

$$\frac{dF}{dt} + \left[\lambda\theta_z + \mu\sum_k \frac{d\theta_{p_k}}{dx_k}\right]F = 0$$

becomes

$$\mathbf{S}\sum_{i=1}^{n} F_{A_i}\left\{\sum_{k=1}^{n} A_k\frac{d\theta_{p_k}}{dx_i} - \sum_{jk} P_{jk}\left(\left(\frac{d^3\theta}{dx_i dx_j dx_k}\right)\right)\right\}$$

$$+ \mathbf{S}F_Z\left\{\sum_{k=1}^{n} A_k\theta_{p_k z} - Z\theta_z - \sum_{k=1}^{n}\frac{d\theta_z}{dx_k}P_k - \sum_{jk} P_{jk}\frac{\partial}{\partial z}\left(\frac{d^2\theta}{dx_j dx_k}\right)\right\}$$

$$+ \mathbf{S}\sum_{i=1}^{m} F_{P_i}\left\{\sum_{r=1}^{n} A_r\theta_{p_i p_r} - \sum_{r=1}^{n} P_r\frac{d\theta_{p_i}}{dx_r} - P_i\theta_z + \sum_{jk} P_{jk}\frac{\partial}{\partial p_i}\left(\frac{d^2\theta}{dx_j dx_k}\right)\right\}$$

$$- \mathbf{S}\sum_{hk} F_{P_{hk}}\left\{\sum_{\alpha\beta} P_{\alpha\beta}\frac{\partial}{\partial p_{hk}}\left(\frac{d^2\theta}{dx_\alpha dx_\beta}\right)\right\} + \left[\lambda\theta_z + \mu\sum_{k=1}^{n}\frac{d\theta_{p_k}}{dx_k}\right]F = 0.$$

Equating coefficients of the derivatives of θ to zero, we obtain the following system of equations:

$$\mathbf{S}\,[\,P_{jk}F_{A_i}+P_{ki}F_{A_j}+P_{ij}F_{A_k}\,]=0 \qquad (i\neq j,\ j\neq k,\ k\neq i), \tag{17}$$

$$\mathbf{S}\,[\,P_{ij}F_{A_i}+P_{ii}F_{A_j}\,]=0, \tag{18}$$

$$\mathbf{S}\,[\,P_{ij}F_Z\,]=0, \tag{19}$$

$$\mathbf{S}\,[\,P_{ij}F_{P_i}\,]=0, \tag{20}$$

$$\mathbf{S}\left[\,A_iF_{A_i}-P_iF_{P_i}-P_{ii}F_{P_{ii}}-\sum_{j=1}^{n}P_{ij}F_{P_{ij}}\right]+\mu F=0, \tag{21}$$

$$\mathbf{S}\left[\,A_iF_{A_j}-P_jF_{P_i}-\sum_{k=1}^{n}P_{kj}F_{P_{ki}}-P_{jj}F_{P_{ji}}\right]=0 \qquad (i\neq j), \tag{22}$$

$$\mathbf{S}\,[\,P_iF_Z\,]=0, \tag{23}$$

$$\mathbf{S}\,[\,A_iF_Z\,]=0, \tag{24}$$

$$\mathbf{S}\,[\,A_jF_{P_i}+A_iF_{P_j}\,]=0, \tag{25}$$

$$\mathbf{S}\left[\,ZF_Z+\sum_{i=1}^{n}P_iF_{P_i}+\sum_{\alpha\beta}P_{\alpha\beta}F_{P_{\alpha\beta}}\right]=\lambda F \qquad (i,j,k=1,2,\cdots,n). \tag{26}$$

There now arise two cases to be considered. In the first case $P_{ij}=0$ $(i,j=1,2,\cdots,n)$, and our expressions are therefore of the first order. In this case the equations become

$$\mathbf{S}\,[\,A_iF_{A_i}-P_iF_{P_i}\,]+\mu F=0,$$

$$\mathbf{S}\,[\,A_iF_{A_j}-P_jF_{P_i}\,]=0,$$

$$\mathbf{S}\,[\,P_iF_Z\,]=0,\ \mathbf{S}\,[\,A_iF_Z\,]=0,$$

$$\mathbf{S}\,[\,A_jF_{P_i}+A_iF_{P_j}\,]=0,$$

$$\mathbf{S}\left[\,ZF_Z+\sum_{i=1}^{n}P_iF_{P_i}\right]=\lambda F \qquad (i,j=1,2,\ldots n).$$

These equations are almost identical with the set (11)...(16); they must of course possess the same integrals as that set, together with others arising from the facts that equations of the type (12) have not now to be satisfied, and that μ, which is an arbitrary constant, has the particular value zero in the first set of equations.

It is easy to see that the integrals still to be found are functions of the variables A alone, and therefore they satisfy the equations

$$\mathbf{S}A_iF_{A_i}+\mu F=0,$$

$$\mathbf{S}A_iF_{A_j}=0 \qquad (i,j=1,2,\cdots,n).$$

Assuming that there are r expressions of the first order whose invariants we are seeking, the solutions of the above equations are easily seen to be the n-row determinants of the matrix,

$$\begin{Vmatrix} A_{1,1} & A_{2,1} & \cdots & A_{r,1} \\ A_{1,2} & & \cdots & \vdots \\ \vdots & & & \\ A_{1,n} & & \cdots & A_{r,n} \end{Vmatrix},$$

provided that r is greater than n.

There are no additional solutions if r is less than n.

These solutions are the Jacobians of sets of n of the forms.

We shall now consider the case in which the quantities P_{ij} are not all zero.

Before discussing the general case, we shall consider the case in which there are only two independent variables.

We shall take in order the cases in which there are one, two, three expressions of the second order whose invariants we are seeking.

In the case of one such expression,

$$\begin{aligned} &F_{A_1} = F_{A_2} = 0 = F_Z = F_{P_1} = F_{P_2}, \\ &2P_{11}F_{P_{11}} + P_{12}F_{P_{12}} = \mu F, \\ &P_{12}F_{P_{12}} + 2P_{22}F_{P_{22}} = \mu F, \\ &2P_{11}F_{P_{12}} + P_{21}F_{P_{22}} = 0, \\ &2P_{22}F_{P_{12}} + P_{21}F_{P_{11}} = 0. \end{aligned}$$

These equations show that F must be a homogeneous function of the algebraic invariant of the binary form

$$(P_{11},\ \tfrac{1}{2}P_{12},\ P_{22}\{\ *\ )^2.$$

Hence $F \equiv$ const. $\times\ (P_{12}^2 - 4P_{11}P_{22})^{\mu/2}$.

To interpret this invariant, suppose that

$$f(x_1\ x_2\ z\ p_1\ p_2\ p_{11}\ p_{12}\ p_{22}) = 0$$

is a differential equation of the second order.

Let $z = \phi(x_1 x_2)$ be some non-singular solution of this equation. We define two directions on this integral surface by means of the equation

$$P_{22}dx_1^2 - P_{12}dx_1dx_2 + P_{11}dx_2^2 = 0.$$

Along one of the curves thus determined on the particular integral surface, x_1, x_2, z, p_1, p_2, p_{11}, p_{12}, p_{22} are functions of a single parameter while dz, dp_1, dp_2 are determined from the equations

$$dz = p_1 dx_1 + p_2 dx_2,$$

$$dp_1 = p_{11} dx_1 + p_{12} dx_2,$$

$$dp_2 = p_{12} dx_1 + p_{22} dx_2.$$

Also

$$\frac{df}{dx_1} + P_{11}\frac{dp_{11}}{dx_1} + P_{22}\frac{dp_{12}}{dx_2} = 0,$$

$$\frac{df}{dx_2} + P_{11}\frac{dp_{12}}{dx_1} + P_{22}\frac{dp_{22}}{dx_2} = 0.$$

The system of seven equations thus obtained are equivalent to six distinct relations, and they determine the "characteristics" of the given equation $f = 0$.

The fundamental property in connection with the curves obtained is that if two integral surfaces have contact of the first order along a characteristic, and if they have contact of the second order at any one point of this curve, they have contact of the second order all along the curve.*

We notice that the directions of the curves are given at every point by means of $z = \phi(x_1 x_2)$ and $P_{11} dx_2^2 - P_{12} dx_1 dx_2 + P_{22} dx_1^2 = 0$.

Now the transformation considered changes an integral surface into an integral surface, and also a characteristic upon an integral surface into a characteristic upon the transformed surface. We therefore expect $P_{11} dx_2^2 - \cdots = 0$ to be an invariant of the expression considered, and we also expect any function geometrically connected with it to be an invariant. $P_{12}^2 - 4P_{11}P_{22}$ was therefore *a priori* to be expected as an invariant.

We next consider the case of two expressions f_1 and f_2.

It is easy to see that the only invariants are the algebraic invariants of the two binary forms

$$(P_{1,11}\tfrac{1}{2}P_{1,12}P_{1,22}\{\ *\ )^2,$$

$$(P_{2,11}\tfrac{1}{2}P_{2,12}P_{2,22}\{\ *\ )^2,$$

unless the condition

$$I \equiv \begin{vmatrix} P_{1,12} & P_{2,12} \\ P_{1,11} & P_{2,11} \end{vmatrix} \times \begin{vmatrix} P_{1,22} & P_{2,22} \\ P_{1,12} & P_{2,12} \end{vmatrix} - \begin{vmatrix} P_{1,11} & P_{2,11} \\ P_{1,22} & P_{2,22} \end{vmatrix}^2 = 0$$

holds.

These invariants have an immediate interpretation from the theory of characteristics. I is itself one of the invariants above mentioned, and $I = 0$ is the condition that the two quadratic forms above mentioned, when equated to zero, have a common root. Therefore, unless the characteristics of $f_1 = 0$ have one direction common with those of $f_2 = 0$ at every point on a common integral surface, the two expressions have only three functionally independent invariants of

* GOURSAT, *Equations aux dérivées partielles du second ordre*, vol. 1, pp. 170 seq.

our type, namely those of two quadratic forms. Suppose now that I is zero. Let m be used to denote the common root mentioned above. Then

$$m = -\frac{P_{1,11}P_{2,22} - P_{2,11}P_{1,22}}{P_{1,12}P_{2,11} - P_{2,12}P_{1,11}} = -\frac{P_{1,22}P_{2,12} - P_{2,22}P_{1,12}}{P_{1,11}P_{2,22} - P_{2,11}P_{1,22}}.$$

Let H denote $A_{1,1}P_{2,11} - A_{2,1}P_{1,11}$ and let K denote $A_{1,2}P_{2,22} - A_{2,2}P_{1,22}$. Then it easily follows that $K + mH$ satisfies our system of equations provided it is zero, and further, this is the only additional solution the system can have.

We may verify that

$$\frac{d}{dt}(K + mH) = \left\{ 3\theta_z - 3\frac{d\theta_{P_1}}{dx_1} + m\frac{d\theta_{P_1}}{dx_2} + \frac{1}{m}\frac{d\theta_{P_2}}{dx_1} - \frac{d\theta_{P_2}}{dx_1} \right\} \times (K + mH).$$

Hence $K + mH$ is not an invariant of our type, although $K + mH = 0$ is an invariant relation.

These two equations

$$I = 0,$$

$$K + mH = 0$$

have an important signification in the theory of differential equations. They are* the conditions that the two equations $f_1 = 0, f_2 = 0$, form a system in involution, in other words, they are the conditions that the two equations have a system of common integrals depending on an infinite number of arbitrary constants.

We shall now consider invariants of three expressions f_1, f_2, f_3.

From the system of equations it follows that $F_Z = 0$ unless all the 3-row determinants of the matrix

$$\begin{vmatrix} A_{1,1} & A_{1,2} & P_{1,1} & P_{1,2} & P_{1,11} & P_{1,12} & P_{1,22} \\ A_{2,1} & A_{2,2} & P_{2,1} & P_{2,2} & P_{2,11} & P_{2,12} & P_{2,22} \\ A_{3,1} & A_{3,2} & P_{3,1} & P_{3,2} & P_{3,11} & P_{3,12} & P_{3,22} \end{vmatrix}$$

are zero.

Also F_{P_λ} is zero unless all the 3-row determinants of the matrix

$$\begin{vmatrix} A_{1,1} & A_{1,2} & P_{1,11} & P_{1,12} & P_{1,22} \\ A_{2,1} & A_{2,2} & P_{2,11} & P_{2,12} & P_{2,22} \\ A_{3,1} & A_{3,2} & P_{3,11} & P_{3,12} & P_{3,22} \end{vmatrix}$$

are zero.

Assuming that the conditions mentioned are not satisfied we see that

$$F_Z = 0, \qquad F_{P_1} = 0, \qquad F_{P_2} = 0.$$

* GOURSAT, *Equations aux dérivées partielles du second ordre*, vol. 2, p. 40 and p. 76.

The functionally independent solutions of our equations are then seen to be the algebraic invariants of the binary forms $(K, H\,|\,*\,)$:

$$(P_{1,11},\ \tfrac{1}{2}P_{1,12},\ P_{1,22}\,|\,*\,)^2,$$
$$(P_{2,11},\ \tfrac{1}{2}P_{2,12},\ P_{2,22}\,|\,*\,)^2,$$
$$(P_{3,11},\ \tfrac{1}{2}P_{3,12},\ P_{3,22}\,|\,*\,)^2,$$

where

$$H \equiv \begin{vmatrix} P_{1,11} & P_{2,11} & P_{3,11} \\ A_{1,1} & A_{2,1} & A_{3,1} \\ P_{1,22} & P_{2,22} & P_{3,22} \end{vmatrix} - \begin{vmatrix} A_{1,2} & \cdots \\ P_{1,12} & \cdots \\ P_{1,22} & \cdots \end{vmatrix},$$

$$K \equiv \begin{vmatrix} P_{1,11} & \cdots \\ A_{1,2} & \cdots \\ P_{1,22} & \cdots \end{vmatrix} - \begin{vmatrix} P_{1,11} & \cdots \\ P_{1,12} & \cdots \\ A_{1,1} & \cdots \end{vmatrix}.$$

It is important to notice the meaning of the equations $H=0$, $K=0$. They are in fact the conditions that the three equations $f_1=0$, $f_2=0$, $f_3=0$ have a common integral surface.*

The three quadratic binary forms and their invariants have as before, immediate interpretation from the theory of characteristics, but the linear form $Kdx_2 - Hdx_1$ has no such immediate interpretation.

As another example of invariants of this type we shall now consider the case in which there are two expressions, one of the second order and the other of the first.

The equations are readily solved, and the solutions are the algebraic invariants of the two forms

$$(P_{1,11},\ \tfrac{1}{2}P_{1,12},\ P_{1,22}\,|\,*\,)^2,$$
$$(A_{2,2},\ -A_{2,1}\,|\,*\,),$$

where $P_{2,11}$, $P_{2,12}$, $P_{2,22}$ are all zero. We know that

$$P_{1,11}dx_2^2 - P_{1,12}dx_1dx_2 + P_{1,22}dx_1^2 = 0$$

is the equation for the directions of the characteristics of $f_1=0$. It seems, therefore, important to consider the meaning of

$$A_{2,2}dx_2 - (-A_{2,1})dx_1,$$

or

$$A_{2,1}dx_1 + A_{2,2}dx_2.$$

But this is equal to df_2, provided

$$dz - p_1dx_1 - p_2dx_2 = 0,\quad dp_1 - p_{11}dx_1 - p_{12}dx_2 = 0,\quad dp_2 - p_{12}dx_1 - p_{22}dx_2 = 0.$$

* VALYI, Crelle's Journal, vol. 95 (1883) p. 100; GOURSAT, loc. cit., vol. 2, p. 199.

We thus have an interpretation of both invariants.

We now return to the general case, when there are n independent variables. The equations (17)—(26) possess all the invariants of our type as solutions.

Suppose that there are r expressions f. Then from (19), (23), (24), $F_Z = 0$ unless r is greater than $\frac{1}{2}n(n+1)+2n$, or unless all the r-row determinants of the matrix

$$\begin{Vmatrix} P_{1,11} & P_{1,12} & \cdots & P_{1,nn} & P_{1,1} & \cdots & P_{1,n} & A_{1,1} & \cdots & A_{1,n} \\ \vdots & \vdots & & & & & & & & \\ P_{r,11} & P_{r,12} & \cdots & & & & & & & A_{r,n} \end{Vmatrix}$$

vanish.

Also the equations (17) (18) show that unless the determinants of another matrix vanish, $F_{A_i} = 0$ for all values of i.

Suppose that F_Z, $F_{A_1}, \cdots, F_{A_n}$ are all zero. There remains the system of equations for F

$$\mathbf{S}(P_{ij} F_{P_i}) = 0,$$

$$\mathbf{S}\left[P_i F_{P_i} + P_{ii} F_{P_{ii}} + \sum_{i=1}^{n} P_{ij} F_{P_{ij}}\right] - \mu F = 0,$$

$$\mathbf{S}\left[P_j F_{P_i} + \sum_{k=1}^{n} P_{kj} F_{P_{ki}} + P_{jj} F_{P_{ji}}\right] = 0,$$

$$\mathbf{S}\left[A_j F_{P_i} + A_i F_{P_j}\right] = 0,$$

$$\mathbf{S}\left[\sum_{i=1}^{n} P_i F_{P_i} + \sum_{\alpha\beta} P_{\alpha\beta} F_{P_{\alpha\beta}}\right] = \lambda F \qquad (i, j, k = 1, \cdots, n),$$

and the fourth of these equations shows that F_{P_i} is zero for all values of i. Hence the system becomes

$$\mathbf{S}\left[P_{ii} F_{P_{ii}} + \sum_{j=1}^{n} P_{ij} F_{P_{ij}}\right] - \mu F = 0,$$

$$\mathbf{S}\left[\sum_{k=1}^{n} P_{kj} F_{P_{ki}} + P_{jj} F_{P_{ji}}\right] = 0,$$

$$\mathbf{S}\left[\sum_{\alpha\beta} P_{\alpha\beta} F_{P_{\alpha\beta}}\right] = \lambda F \qquad (i, j, k = 1, \cdots, n).$$

Hence $\lambda = n\mu/2$, and the invariants required are the algebraic invariants of the system of quadratic forms

$$\sum_{\alpha\beta} P_{\alpha\beta} x_\alpha x_\beta,$$

where none of the magnitudes $P_{\alpha\beta}$ is repeated and $P_{\alpha\beta} = P_{\beta\alpha}$.

These quadratic forms are easily seen to be those which FORSYTH * calls

* FORSYTH, Philosophical Transactions, ser. A, vol. 191 (1898), p. 2.

"characteristic invariants" when there are only three independent variables. This name might with advantage be extended to the general case in which there are n independent variables.

We now return to the case in which $F_{A_1} F_{A_2}$, etc., are not all zero.

Suppose that there are three independent variables, and two expressions f. Then if F_{A_1}, etc., are not zero, all the 6-row determinants of the matrix

$$\begin{vmatrix}
P_{1,11} & 0 & 0 & P_{1,12} & P_{1,22} & 0 & 0 & P_{1,31} & P_{1,33} & P_{1,23} \\
P_{2,11} & 0 & 0 & P_{2,12} & P_{2,22} & 0 & 0 & P_{2,31} & P_{2,33} & P_{2,23} \\
0 & P_{1,22} & 0 & P_{1,11} & P_{1,12} & P_{1,23} & P_{1,33} & 0 & 0 & P_{1,31} \\
0 & P_{2,22} & 0 & P_{2,11} & P_{2,12} & P_{2,23} & P_{2,33} & 0 & 0 & P_{2,31} \\
0 & 0 & P_{1,33} & 0 & 0 & P_{1,22} & P_{1,23} & P_{1,11} & P_{1,31} & P_{1,12} \\
0 & 0 & P_{2,33} & 0 & 0 & P_{2,22} & P_{2,23} & P_{2,11} & P_{2,31} & P_{2,12}
\end{vmatrix}$$

must vanish.

Let S_i denote the characteristic invariant of f_i, then it is easily seen that if we construct the cubic forms

$$S_1 x_1,\ S_1 x_2,\ S_1 x_3,\ S_2 x_1,\ S_2 x_2,\ S_2 x_3,$$

the above matrix is the matrix of the coefficients.

Hence, if our conditions hold, the above six cubics must belong to a five fold linear system. Expressing this condition we see that $S_1 L_1 + S_2 L_2 \equiv 0$, where L_1 and L_2 are certain linear forms.

Hence either S_1 and S_2 both break up into linear factors, or S_2 is equivalent to S_1.

In the case in which there are three expressions f, the conditions give

$$S_1 L_1 + S_2 L_2 + S_3 L_3 \equiv 0,$$

where L_1, L_2, L_3 are linear. Hence, in general S_1, S_2, S_3, regarded as conics, have two common points.

The generalization is immediate, and the condition in order that invariants involving the magnitudes A, of r expressions f in n independent variables exist, are equivalent to the conditions that r linear forms $L_1, \cdots, L_r$ should exist such that

$$\sum_{i=1}^{r} S_i L_i \equiv 0$$

identically, when S_i is the characteristic invariant of f_i. It is readily seen that these conditions may be expressed by the vanishing of certain algebraic invariants of the r quadratic forms S.

An upper limit to the number of these conditions may readily be obtained.

This upper limit is $n(n+1)(n+2)/3! - nr + 1$. The number may fall below this in certain cases, for example, if $n = 3$ and $r = 2$ it is 4, whilst

$$\frac{n(n+1)(n+2)}{3!} - nr + 1 = 5.$$

If all the linear expressions L are equivalent, the conditions require that constants λ can be found such that $\sum \lambda S = 0$.

The number of conditions for this is readily seen to be

$$\frac{n(n+1)}{2} - r + 1.$$

This number is less than the previous one if $[n(n+1)/6 - r](n-1) > 0$, and the second conditions are all independent.

Hence, if r is $< \frac{1}{6}n(n+1)$ the imposition of $n(n+1)/2 - r + 1$ conditions is sufficient to cause the remainder to be satisfied.

Further consideration of this question will be omitted from the present paper.

Suppose the conditions in question to be satisfied. We then obtain a solution of the set of equations in F_A which is a determinant linear in the magnitudes A. Call this determinant Δ, then in a manner strictly analogous to the case when $n = 2$, it may be shown that $\Delta = 0$ is an invariant relation, provided the previous set of conditions holds.

If there are solutions of the system of equations considered which involve the magnitudes A, the derivatives F_{P_i} are not necessarily zero. From the equations of type (20), we see that, if $F_{P_i} \neq 0$, the matrix

$$\begin{vmatrix} P_{1,11} & \cdots & P_{1,ij} & \cdots & P_{1,nn} \\ \vdots & & \vdots & & \\ P_{r,11} & \cdots & P_{r,ij} & \cdots & P_{r,nn} \end{vmatrix}$$

must have all its r-row determinants zero.

In addition, all the equations of type (25),

$$\mathfrak{S}(A_j F_{P_i} + A_i F_{P_j}) = 0,$$

must be satisfied.

Now take any one of the determinants of the above matrix, replace one of its columns by the magnitudes $P_{1,i}, P_{2,i}, \cdots, P_{r,i}$. Call the determinant then formed Δ_i. Call the similar determinant with A_i instead of P_i, M_i. Then P_i only enters through Δ_i, and the equations (25) become

$$M_j F_{\Delta_i} + M_i F_{\Delta_j} = 0 \qquad (i, j = 1, \cdots, r).$$

Hence if an invariant contains P_i, we see from the case when $i = j$, that M_i must be zero. If $M_i = 0$ and $F_{\Delta_i} \neq 0$, we see that $M_j = 0$ $(j = 1, \cdots, r)$; and if $F_{\Delta_i} = 0$ and $M_i \neq 0$, we see that $F_{\Delta_j} = 0$ $(j = 1, 2, \cdots, r)$.

Hence either none of the magnitudes $P_i (i = 1, \cdots, r)$ occur in any invariant, or all the magnitudes $M_i (i = 1, \cdots, r)$ are zero.

It is easy to see that if all the given conditions are satisfied, then the magnitudes Δ_i satisfy the remaining equations, and therefore these magnitudes Δ_i are invariants.

There is no invariant involving Z unless all the conditions given in connection with the magnitude P_i are satisfied and, in addition,

$$\Delta_1 = \Delta_2 = \cdots = \Delta_r = 0.$$

If all these conditions are satisfied, there is an invariant involving Z given by replacing any column in any r-row determinant of the matrix

$$\left\| \begin{matrix} P_{1,11} & \cdots & P_{1,ij} & \cdots & P_{1,nn} \\ \vdots & & \vdots & & \vdots \\ P_{r,11} & \cdots & P_{r,ij} & \cdots & P_{r,nn} \end{matrix} \right\|$$

by $Z_1, Z_2, \cdots, Z_r$.

§ 3.

We have not as yet considered invariants which involve the magnitudes $dx_i, dz, dp_i, dp_{ij} (i, j = 1, \cdots, n)$.

It is clear that invariants of this type do exist. For example, it is easy to verify that

$$\sum_{\alpha, \beta} P_{\alpha\beta}\, dx_\alpha dx_\beta$$

is such an invariant.

The work is somewhat simplified if we take as variables

$$\begin{aligned} dx_i &\equiv a_i, \\ dz - \sum_i p_i dx_i &\equiv u, \\ dp_i - \sum_j p_{ij} dx_j &\equiv v_i, \\ dp_{ij} &\equiv c_{ij}. \end{aligned} \qquad (i, j = 1, \cdots, n).$$

The increments of these magnitudes are readily obtained. We have

$$-\frac{da_i}{dt} = \sum_k \frac{d\theta_{p_i}}{dx_k} a_k + \frac{\partial \theta_{p_i}}{\partial z} u + \sum_k \frac{\partial^2 \theta}{\partial p_i \partial p_k} v_k, \qquad \frac{du}{dt} = \theta_z u,$$

$$\frac{dv_i}{dt} = \frac{d\theta_z}{dx_i} u + \sum_k \frac{d\theta_{p_k}}{dx_i} v_k + \theta_z v_i,$$

$$\frac{dc_{ij}}{dt} = \sum_k \left(\left(\frac{d^3\theta}{dx_i dx_j dx_k} \right) \right) a_k + \left(\frac{d^2\theta_z}{dx_i dx_j} \right) u$$

$$+ \sum_k \left(\frac{d^2\theta_{p_k}}{dx_i dx_j} \right) v_k + \frac{d\theta_z}{dx_j} v_i + \frac{d\theta_z}{dx_i} v_j + \sum_{\alpha\beta} \frac{\partial}{\partial p_{\alpha\beta}} \left(\frac{d^2\theta}{dx_i dx_j} \right) c_{\alpha\beta}.$$

If F is an invariant, the equation which it satisfies is similar to the one given in the previous section, but it contains the additional terms

$$-\sum_i F_{a_i}\left\{\sum_k \frac{d\theta_{P_i}}{dx_k}a_k + \frac{\partial\theta_{P_i}}{\partial z}u + \sum_k \frac{\partial^2\theta}{\partial p_i \partial p_k}v_k\right\}$$

$$+ F_u\theta_z u + F_{v_i}\left\{\frac{d\theta_z}{dx_i}u + \theta_z v_i + \sum_k \frac{d\theta_{P_k}}{dx_i}v_k\right\}$$

$$+\sum_{ij}\left\{\sum_k\left(\left(\frac{d^3\theta}{dx_i dx_j dx_k}\right)\right)a_k + \left(\frac{d^2\theta_z}{dx_i dx_j}\right)u\right.$$

$$\left.+\sum_k\left(\frac{d^2\theta_{P_k}}{dx_i dx_j}\right)v_k + \frac{d\theta_z}{dx_j}v_i + \frac{d\theta_z}{dx_i}v_j + \sum_{\alpha\beta}\frac{\partial}{\partial p_{\alpha\beta}}\left(\frac{d^2\theta}{dx_i dx_j}\right)c_{\alpha\beta}\right\}F_{c_{ij}}.$$

The equations for F are now

$$\mathbf{S}[P_{jk}F_{A_i} + P_{ki}F_{A_j} + P_{ij}F_{A_k}] - a_k F_{c_{ij}} - a_i F_{c_{jk}} - a_j F_{c_{ik}} = 0 \qquad (i \neq j, j \neq k, k \neq i), \tag{27}$$

$$\mathbf{S}[P_{ij}F_{A_i} + P_{ii}F_{A_j}] - a_j F_{c_{ii}} - a_i F_{c_{ij}} = 0, \tag{28}$$

$$\mathbf{S}P_{ij}F_Z - uF_{c_{ij}} = 0, \tag{29}$$

$$\mathbf{S}P_{ij}F_{P_k} - v_k F_{c_{ki}} = 0, \tag{30}$$

$$\mathbf{S}\left[A_i F_{A_i} - P_i F_{P_i} - P_{ii}F_{P_{ii}} - \sum_{j=1}^{n} P_{ij}F_{P_{ij}}\right] + \mu F$$

$$- a_i F_{a_i} + v_i F_{v_i} + c_{ii}F_{c_{ii}} + \sum_{j=1}^{m} c_{ij}F_{c_{ij}} = 0, \tag{31}$$

$$\mathbf{S}\left[A_i F_{A_j} - P_j F_{P_i} - \sum_{k=1}^{n} P_{kj}F_{P_{ki}} - P_{jj}F_{P_{ji}}\right]$$

$$- a_j F_{a_i} + v_i F_{v_j} + c_{ij}F_{c_{jj}} + \sum_{k=1}^{m} c_{ki}F_{c_{kj}} = 0, \tag{32}$$

$$\mathbf{S}[P_i F_Z] - uF_{v_i} = 0, \tag{33}$$

$$\mathbf{S}A_i F_Z - uF_{a_i} = 0, \tag{34}$$

$$\mathbf{S}[A_j F_{P_i} + A_i F_{P_j}] - v_j F_{a_i} - v_i F_{a_j} = 0, \tag{35}$$

$$\mathbf{S}\left[ZF_Z + \sum_{i=1}^{n} P_i F_{P_i} + \sum_{\alpha\beta} P_{\alpha\beta}F_{P_{\alpha\beta}}\right] - \left(\sum_{\alpha\beta} c_{\alpha\beta}F_{c_{\alpha\beta}} + uF_u + \sum_{i=1}^{n} v_i F_{v_i}\right)$$

$$= \lambda F \qquad (i, j, k, \alpha, \beta, = 1, \cdots, n). \tag{36}$$

We first consider the particular case when all the original expressions f are of the first order.

In this case $P_{ij} = 0\,(i, j = 1, \cdots, n)$, and it may be readily seen that F does not in general involve any of the magnitudes $c_{\alpha\beta}$.

We have therefore the reduced system of equations for F:

$$\mathbf{S}\,[A_i F_{A_i} - P_i F_{P_i}] + uF - a_i F_{a_i} + v_i F_{v_i} = 0, \tag{37}$$

$$\mathbf{S}\,[A_i F_{A_j} - P_j F_{P_i}] + v_i F_{v_j} - a_j F_{a_i} = 0, \tag{38}$$

$$\mathbf{S} P_i F_Z - u F_{v_i} = 0, \tag{39}$$

$$\mathbf{S} A_i F_Z - u F_{a_i} = 0, \tag{40}$$

$$\mathbf{S}\,[A_j F_{P_i} + A_i F_{P_j}] - v_j F_{a_i} - v_i F_{a_j} = 0, \tag{41}$$

$$\mathbf{S}\left[ZF_Z + \sum_{i=1}^{n} P_i F_{P_i}\right] - uF_u - \sum_{i=1}^{n} v_i F_{v_i} = \lambda F \qquad (i, j = 1, \cdots, n). \tag{42}$$

From (39) and (40) we deduce that F must be a function of u, W, P_i, A_i, where

$$W = uZ + \sum a_i A_i + \sum v_i P_i;$$

and the equations for F are now

$$\mathbf{S}\,[A_i F_{A_i} - P_i F_{P_i}] + \mu F = 0, \tag{43}$$

$$\mathbf{S}\,[A_i F_{A_j} - P_j F_{P_i}] = 0, \tag{44}$$

$$\mathbf{S}\,[A_j F_{P_i} + A_i F_{P_j}] = 0, \tag{45}$$

$$\mathbf{S}\left[\sum_{i=1}^{n} P_i F_{P_i}\right] + uF_u = \lambda F \qquad (i, j, = 1, \cdots, n). \tag{46}$$

Hence the quantities W are absolute invariants, and in addition there are the invariants given in the previous section. Also u is an invariant.

The additional invariants obtained are readily interpreted. The function

$$u \equiv dz - \sum p_i dx_i,$$

is an invariant arising in connection with the contact transformation itself, and W may easily be shown to be

$$df \equiv \frac{\partial f}{\partial z} dz + \sum_i \frac{\partial f}{\partial x_i} dx_i + \sum_i \frac{\partial f}{\partial p_i} dp_i.$$

Suppose that $u = 0$. Then unless the r-row determinants of the matrix

$$\begin{vmatrix} A_{1,1} & \cdots & A_{1,n} & P_{1,1} & \cdots & P_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{r,1} & & A_{r,n} & P_{r,n} & \cdots & P_{r,n} \end{vmatrix}$$

are all zero, and F_Z is zero.

If we write $v_i = A_{r+1,i}$ and $a_i = -P_{r+1,i}$, the set of equations becomes the same as that given previously in which there were $(r+1)$ expressions and in

which the variables dz, dx_i, dp_i, did not occur, the last equation only being different. The additional solutions are therefore $[f_i, f_{r+1}]$ $(i=1, \cdots, r)$, which are $df_1, df_2, \cdots, df_r$ subject to the condition that $u = 0$.

We now consider expressions F of the second order. Suppose first that $u \neq 0$ and that $v_1, \cdots, v_n$ are not all zero.

Write L for

$$uZ + \sum_i a_i A_i + \sum_i v_i P_i + \sum_{\alpha\beta} c_{\alpha\beta} P_{\alpha\beta},$$

then F is readily shown to be a function of the variables u, L, v_i, P_{ij}, A_i $(i, j = 1, \cdots, n)$, which satisfies the system of equations

$$\mathbf{S}[P_{jk} F_{A_i} + P_{ki} F_{A_j} + P_{ij} F_{A_k}] = 0 \quad (j \neq k, k \neq i, i \neq j),$$

$$\mathbf{S}[P_{ij} F_{A_i} + P_{ii} F_{A_j}] = 0,$$

$$\mathbf{S}\left[A_i F_{A_i} - P_{ii} F_{P_{ii}} - \sum_{j=1}^{n} P_{ij} F_{P_{ij}}\right] + \mu F + v_i F_{v_i} = 0,$$

$$\mathbf{S}\left[A_i F_{A_j} - P_{jj} F_{P_{ji}} - \sum_{k=1}^{n} P_{kj} F_{P_{ki}}\right] + v_i F_{v_j} = 0,$$

$$\mathbf{S}\left(\sum_{\alpha\beta} P_{\alpha\beta} F_{P_{\alpha\beta}}\right) + u F_u + \sum_{i=1}^{n} v_i F_{v_i} = \lambda F.$$

The magnitudes L are therefore invariants. In addition u is an invariant, and the remaining invariants are solutions of the given system.

L is readily shown to be

$$df \equiv \sum_i \frac{\partial f}{\partial x_i} dx_i + \frac{\partial f}{\partial z} dz + \sum_i \frac{\partial f}{\partial p_i} dp_i + \sum_{\alpha\beta} \frac{\partial f}{\partial p_{\alpha\beta}} dp_{\alpha\beta},$$

and it therefore admits of an immediate interpretation. There are also the invariants obtained in the preceding section which do not involve the magnitudes v_i $(i = 1, \cdots, n)$ and the remaining invariants are those solutions of the set of equations last given which involve these magnitudes.

We see that when certain conditions given in the preceding section hold, the quantities F_{A_i} $(i = 1, \cdots, n)$ are all zero.

There remains the system

$$\mathbf{S}\left(P_{ii} F_{P_{ii}} + \sum_{j=1}^{n} P_{ij} F_{P_{ij}}\right) - v_i F_{v_i} = \mu F,$$

$$\mathbf{S}\left(P_{jj} F_{P_{ji}} + \sum_{k=1}^{n} P_{kj} F_{P_{ki}}\right) - v_i F_{v_j} = 0,$$

$$\mathbf{S}\left(\sum_{\alpha\beta} P_{\alpha\beta} F_{P_{\alpha\beta}}\right) + \sum_{i=1}^{n} v_i F_{v_i} = \lambda F \quad (i, j, \alpha, \beta, = 1, \cdots, n).$$

The solution of these equations is a function of the algebraic invariants and covariants of the r quadratic forms

$$\sum_{ij} P_{ij} v_i v_j .$$

In addition, this function must be homogeneous in the variables v.

These forms

$$\sum_{ij} P_{ij} v_i v_j$$

are those which have been referred to earlier as the Characteristic Invariants of the expressions f.

Now suppose that u and the quantities v are all zero.

In this case F_z and $F_{P_i}\,(i = 1, \cdots, n)$ are all zero unless all the r-row determinants of the matrix

$$\left\| \begin{matrix} P_{1,11} \cdots P_{1,ij} \cdots P_{1,nn} \\ \vdots \quad\quad \vdots \quad\quad \vdots \\ P_{r,11} \cdots P_{r,ij} \cdots P_{r,nn} \end{matrix} \right\|$$

vanish.

We assume that these conditions are not satisfied, and therefore F is a function of a_i, A_i, P_{ij}, $c_{ij}\,(i, j = 1, \cdots, n)$.

The equations of types (27) and (28) are satisfied by the magnitudes

$$L \equiv \sum_i A_i a_i + \sum_{jk} P_{jk} c_{jk} .$$

It may readily be shown that these are the only independent solutions of this set of equations unless it is possible to make

$$\sum_{i=1}^{r} S_i B_i + AK = 0$$

identically, where any n variables $y_1, y_2, \cdots, y_n$ are taken, and the magnitudes B_i are linear functions of these variables, K is a quadratic function of them, and

$$S_i \equiv \sum_{\alpha\beta} P_{i,\alpha\beta} y_\alpha y_\beta ,$$

$$A \equiv \sum_k a_k y_k .$$

This equation can be satisfied by making the linear functions all equal to A, multiplied by certain constant factors. This leads to solutions of our system of the type L.

If we assume that there is no other way of making the expression considered an identity, we see that the quantities L are the only invariants involving the variables $A_k (k = 1, \cdots, n)$.

It is readily seen that these expressions L are equivalent to the expressions df, subject to the conditions

$$u = 0, \qquad v_k = 0 \qquad (k = 1, 2, \cdots, n).$$

It is easy to show that the remaining solutions of the system are the algebraic invariants of the system of quadratic forms S_i and of the linear form A.

These forms S are again the characteristic invariants of the expressions f; they have however, in this case, the magnitudes y for independent variables.

Combining our results we have the following theorem:

All invariants, of the restricted type considered of r second order expressions f in one dependent and n independent variables are

(1) *Expressions of the type df.*

(2) *Algebraic invariants and covariants of the quadratic forms*

$$\sum_{ij} \frac{\partial f}{\partial p_{ij}} v_i v_j,$$

where the v's are the variables, and

$$v_k = dp_k - \sum_{i=1}^{n} p_{ki} dx_i,$$

provided

$$dz \neq \sum_{i=1}^{n} p_i dx_i,$$

and the v's are not zero.

(3) *Algebraic invariants of the quadratic forms*

$$\sum_{ij} \frac{\partial f}{\partial p_{ij}} y_i y_j,$$

and of the linear form

$$\sum_{i=1}^{n} y_i dx_i,$$

where the y's are the variables, provided that

$$dz = \sum_{i=1}^{n} p_i dx_i, \qquad dp_i = \sum_{k=1}^{n} p_{ik} dx_k.$$

In the above there are two sets of restrictions on the expressions f.

(1) The r-row determinants must not all be zero in the matrix

$$\begin{Vmatrix} P_{1,11} & \cdots & P_{1,ij} & \cdots & P_{1,nn} \\ \vdots & & & & \\ P_{r,11} & \cdots & P_{r,ij} & \cdots & P_{r,nn} \end{Vmatrix}.$$

(2) It must not be possible to satisfy the identity

$$\sum_{i=1}^{r} S_i B_i + AK \equiv 0,$$

where S, B, A, K, are as previously defined, except by making the B's all constant multiples of A.

The case when the second of the above restrictions is removed requires further consideration. It is obvious that it must be removed if r is great enough, but the whole question will be left for future consideration. At present we content ourselves with a discussion of the case when n is two.

For one expression f, $SB + AK \equiv 0$, provided that S admits A as a factor, since we are neglecting the possibility of B being λA, where λ is a constant.

This gives $P_{11}a_2^2 - P_{12}a_1a_2 + P_{22}a_1^2 = 0$, as the condition for the existence of further integrals.

If this condition is satisfied, the equations of type (27) and (28) possess the two solutions

$$\alpha_1 = A_1a_1a_2 + P_{22}a_1c_{12} + P_{11}a_2c_{11},$$

$$\alpha_2 = A_2a_1a_2 + P_{11}a_2c_{12} + P_{22}a_1c_{11},$$

instead of the single one already given.

It may readily be shown that there are no new invariants, but if

$$I \equiv P_{11}a_2^2 - P_{12}a_1a_2 + P_{22}a_1^2 = 0,$$

then the two equations

$$\alpha_1 \equiv A_1a_1a_2 + P_{22}a_1c_{12} + P_{11}a_2c_{11} = 0,$$

$$\alpha_2 = A_2a_1a_2 + P_{11}a_2c_{12} + P_{22}a_1c_{22} = 0$$

are an invariant system.

We observe that I is an invariant.

The equations $I = 0$, $\alpha_1 = 0$, $\alpha_2 = 0$, taken in conjunction with

$$dz = p_1dx_1 + p_2dx_2, \qquad dp_1 = p_{11}dx_1 + p_{12}dx_2, \qquad dp_2 = p_{12}dx_1 + p_{22}dx_2,$$

have an immediate and important interpretation in connection with the differential equation $f = 0$. They are precisely the equations for the characteristics of this differential equation.*

Now suppose that there are two expressions f. In this case there are four equations in seven variables of the types (27) and (28). They therefore possess in general three integrals. Two of these are already known, and are df_1, df_2.

The remaining one may be expressed as the determinant

$$\begin{vmatrix} \lambda_{1,1}, & \lambda_{1,2}, & 0 \\ K_{1,1}, & K_{1,2}, & I_1 \\ K_{1,2}, & K_{2,2}, & I_2 \end{vmatrix} \equiv \Delta,$$

* GOURSAT, *Equations aux dérivées partielles du second ordre*, vol. 2, p. 174.

where

$$\lambda_1 = 2P_{11}a_2 - P_{12}a_1,$$
$$\lambda_2 = 2P_{22}a_1 - P_{12}a_2,$$
$$K_1 = a_1a_2A_2 + P_{11}a_2c_{12} + P_{22}a_1c_{22},$$
$$K_2 = a_1a_2A_1 + P_{22}a_1c_{12} + P_{11}a_2c_{11},$$
$$I = P_{22}a_1^2 - P_{12}a_1a_2 + P_{11}a_2^2.$$

It may easily be shown that Δ/a_1a_2 satisfies the remaining system of equations, and therefore the complete system of integrals is df_1, df_2, Δ/a_1a_2, and the invariants of the binary forms

$$(P_{1,11}, \tfrac{1}{2}P_{1,12}, P_{1,22}\{ *)^2, \qquad (P_{2,11}, \tfrac{1}{2}P_{2,12}, P_{2,22}\{ *)^2, \qquad (a_1, a_2\{ *).$$

We omit for the present the interpretation of Δ, and proceed to consider the case in which there are three expressions f.

In this case the equations of types (27) and (28) have the five functionally independent integrals df_1, df_2, df_3 and

$$H \equiv \begin{vmatrix} P_{1,11} & A_{1,1} & P_{1,22} \\ P_{2,11} & A_{2,1} & P_{2,22} \\ P_{3,11} & A_{3,1} & P_{3,22} \end{vmatrix} - \begin{vmatrix} A_{1,2} & P_{1,12} & P_{1,22} \\ A_{2,2} & P_{2,12} & P_{2,22} \\ A_{3,2} & P_{3,12} & P_{3,22} \end{vmatrix},$$

$$K \equiv \begin{vmatrix} P_{1,11} & A_{1,2} & P_{1,22} \\ P_{2,11} & A_{2,2} & P_{2,22} \\ P_{3,11} & A_{3,2} & P_{3,22} \end{vmatrix} - \begin{vmatrix} P_{1,11} & P_{1,12} & A_{1,1} \\ P_{2,11} & P_{2,12} & A_{2,1} \\ P_{3,11} & P_{3,12} & A_{3,1} \end{vmatrix}.$$

We substitute these integrals in the remaining equations, and the complete system of independent integrals of the equations thus obtained is the system of invariants of the binary forms

$$(P_{1,11}, \tfrac{1}{2}P_{1,12}, P_{1,22}\{ *)^2,$$
$$(P_{2,11}, \tfrac{1}{2}P_{2,12}, P_{2,22}\{ *)^2,$$
$$(P_{3,11}, \tfrac{1}{2}P_{3,12}, P_{3,22}\{ *)^2,$$
$$(K, H\{ *),$$
$$(a_1, a_2\{ *).$$

If we take the variables to be $a_2, -a_1$, we see that the solutions in question are the invariants and covariants of the binary forms

$$Ka_2 - Ha_1,$$
$$P_{1,11}a_2^2 - P_{1,12}a_1a_2 + P_{1,22}a_1^2,$$
$$P_{2,11}a_2^2 - P_{2,12}a_1a_2 + P_{2,22}a_1^2,$$
$$P_{3,11}a_2^2 - P_{3,12}a_1a_2 + P_{3,22}a_1^2.$$

In addition to these there are the three solutions df_1, df_2, df_3, and these are all the functionally independent integrals of our type.

§4.

We shall consider to a small extent the more general type of expression f, that is to say the type in which there are more dependent variables than one.

Let there be m dependent variables, $z_1, \cdots, z_m$, and as before let there be n independent variables, $x_1, \cdots, x_n$. We use the same notation as on pp. 288, 289, and in addition $df/dx_\mu = A_\mu$.

In this case, the most general contact transformation possible may easily be shown to be a point transformation.

Let the expressions f_λ be of the first order, that is to say, let them be functions of the variables x_k, z_i, p_k^i $(i = 1, 2, \cdots, m;\ k = 1, 2, \cdots, n)$.

If F is any first order invariant we have

$$\frac{dF}{dt} = \left(\mu_0 \frac{d\Omega_0}{dt} + \mu_1 \frac{d\Omega_1}{dt}\right) F,$$

where

$$\frac{d\Omega_0}{dt} = \sum_{k=1}^{n} \frac{\partial \xi_k}{\partial x_k} + \sum_{l=1}^{m} \frac{\partial \zeta_i}{\partial z_i},$$

$$\frac{d\Omega_1}{dt} = \sum_{i=1}^{m} \sum_{k=1}^{n} \frac{\partial \pi_k^i}{\partial p_k^i},$$

and μ_0, μ_1, have the meanings given on p. 288.

Expanding dF/dt and equating to zero the coefficients of ξ_k, ζ_i, $\partial\zeta_i/\partial x_\mu$, we see that the variables x, z, do not occur explicitly, and that the variables p, x, only occur through the variables A.

We assume F to be a function of the variables A, Z, P, and

$$a_k \equiv dx_k,\ v_i \equiv dz_i - \sum_{k=1}^{n} p_k^i dx_k \qquad (k=1, \cdots, n,\ i=1, \cdots, m).$$

The values of the various increments involved are the following:

$$\frac{d}{dt} a_k = \sum_{\lambda=1}^{m} \frac{\partial \xi_k}{\partial z_\lambda} v_\lambda + \sum_{\mu=1}^{n} \frac{d\xi_k}{dx_\mu} a_\mu,$$

$$\frac{d}{dt} v_i = \sum_{\lambda=1}^{m} \frac{\partial \theta_i}{\partial z_\lambda} v_\lambda,$$

$$-\frac{dA_k}{dt} = \sum_{r=1}^{n} \frac{d\xi_r}{dx_k} A_r + \sum_{i=1}^{m} \sum_{\mu=1}^{n} P_\mu^i \frac{d^2\theta_i}{dx_\mu dx_k},$$

$$-\frac{dZ_\lambda}{dt} = \sum_{r=1}^{n} \frac{\partial \xi_r}{\partial z_\lambda} A_r + \sum_{s=1}^{m} \frac{\partial \theta_s}{\partial z_\lambda} Z_s + \sum_{i=1}^{m} \sum_{k=1}^{n} P_k^i \frac{d}{dx_k}\left(\frac{\partial \theta_i}{\partial z_\lambda}\right),$$

$$-\frac{dP^h_k}{dt} = \sum_{i=1}^{m} P^i_k \frac{\partial \theta_i}{\partial z_h} - \sum_{\sigma=1}^{n} P^h_\sigma \frac{d\xi_k}{dx_\sigma} \qquad (h, \lambda = 1, 2, \cdots, m;\ k = 1, 2, \cdots, n).$$

Also

$$\mu_0 \frac{d\Omega_0}{dt} + \mu_1 \frac{d\Omega_1}{dt} = \mu \sum_{k=1}^{n} \frac{d\xi_k}{dx_k} + \nu \sum_{h=1}^{m} \frac{\partial \theta_h}{\partial z_h},$$

where μ and ν are constants depending on μ_0 and μ_1.

We substitute these values in the equation for F and equate the coefficients of the various derivatives of ξ and ζ to zero. We thus obtain the following system of equations for F:

$$a_\sigma F_{a_k} - \mathbf{S} A_k F_{A_\sigma} + \mathbf{S} \sum_{i=1}^{m} P^i_\sigma F_{P^i_k} = 0 \qquad (\sigma \neq k), \tag{47}$$

$$a_k F_{a_k} - \mathbf{S} A_k F_{A_k} + \mathbf{S} \sum_{i=1}^{m} P^i_k F_{P^i_k} = \mu F, \tag{48}$$

$$v_\lambda F_{a_k} - \mathbf{S} A_k F_{Z_\lambda} = 0, \tag{49}$$

$$v_\lambda F_{v_i} - \mathbf{S} Z_i F_{Z_\lambda} - \mathbf{S} \sum_{k=1}^{n} P^i_k F_{P^\lambda_k} = 0 \qquad (\lambda \neq i), \tag{50}$$

$$v_i F_{v_i} - \mathbf{S} Z_i F_{Z_i} - \mathbf{S} \sum_{k=1}^{n} P^i_k F_{P^i_k} = \nu F, \tag{51}$$

$$-\mathbf{S} P^i_\mu F_{A_k} - \mathbf{S} P^i_k F_{A_\mu} = 0, \tag{52}$$

$$-\mathbf{S} P^i_k F_{Z_\lambda} = 0 \tag{53}$$

$$(\sigma, k, \mu = 1, \cdots, n);\ (i, \lambda = 1, \cdots, m).$$

The equations (49) give solutions of type

$$\Delta \equiv \sum_{k=1}^{n} a_k A_k + \sum_{i=1}^{m} v_i Z_i,$$

and the equations (53) then show that the Δ's cannot enter into F unless the number r of expressions f_λ is greater than mn.

From equations (52) we deduce that the variables A do not occur in F unless

$$r > \tfrac{1}{2} m(n+1).$$

In the case when m is unity, the equations are not all independent, and r need not satisfy this last condition.

Suppose that r is less than this number, then F is a function of the variables $v_1, \cdots v_m, P^1_1, \cdots P^m_n$, which satisfies the equations

$$\mathbf{S} \sum_{i=1}^{m} P^i_\sigma F_{P^i_k} = 0, \qquad (\sigma \neq k),$$

$$\mathbf{S}\sum_{i=1}^{m} P_k^i F_{P_k^i} = \mu F,$$

$$\mathbf{S}\sum_{k=1}^{n} P_k^i F_{P_k^\lambda} - v_\lambda F_{v_i} = 0,$$

$$\mathbf{S}\sum_{k=1}^{n} P_k^i F_{P_k^i} - v_i F_{v_i} = -\nu F.$$

The last two equations show that F must be an invariant or covariant of the linear forms

$$\sum_{i=1}^{m} P_{\lambda,k}^i v_i \qquad (k=1, 2, \cdots, n\,;\ \lambda=1, 2, \cdots, r).$$

The two first equations show that F must at the same time be an invariant of the linear forms

$$\sum_{k=1}^{n} P_{\lambda,k}^i y_i \qquad (i=1, 2, \cdots, m\,;\ \lambda=1, 2, \cdots, r).$$

BRYN MAWR COLLEGE.

GROUPS OF ORDER p^m, WHICH CONTAIN CYCLIC SUBGROUPS OF ORDER p^{m-3} *

BY

LEWIS IRVING NEIKIRK

Introduction.

The groups of order p^m, which contain self-conjugate cyclic subgroups of orders p^{m-1}, and p^{m-2} respectively, have been determined by BURNSIDE,† and the number of groups of order p^m, which contain cyclic non-self-conjugate subgroups of order p^{m-2} has been given by MILLER.‡

Although in the present state of the theory, the actual tabulation of all groups of order p^m is impracticable, it is of importance to carry the tabulation as far as may be possible. In this paper *all groups of order* p^m (p being an odd prime) *which contain cyclic subgroups of order* p^{m-3} *and none of higher order* are determined. The method of treatment used is entirely abstract in character and, in virtue of its nature, it is possible in each case to give explicitly the generational equations of these groups. They are divided into three classes, and it will be shown that these classes correspond to the three partitions: $(m-3, 3)$, $(m-3, 2, 1)$ and $(m-3, 1, 1, 1)$, of m.

We denote by G an abstract group G of order p^m containing operators of order p^{m-3} and no operator of order greater than p^{m-3}. Let P denote one of these operators of G of order p^{m-3}. The p^3 power of every operator in G is contained in the cyclic subgroup $\{P\}$, otherwise G would be of order greater than p^m. The complete division into classes is effected by the following assumptions:

I. There is in G at least one operator Q_1, such that $Q_1^{p^2}$ is not contained in $\{P\}$.

II. The p^2 power of every operator in G is contained in $\{P\}$, and there is at least one operator Q_1, such that Q_1^p is not contained in $\{P\}$.

III. The pth power of every operator in G is contained in $\{P\}$.

The number of groups for Class I, Class II, and Class III, together with the total number, are given in the table below:

* Presented to the Society April, 25, 1903. Received for publication August 16, 1904, and February 25, 1905.

† *Theory of Groups of a Finite Order*, pp. 75–81.

‡ Transactions, vol. 2 (1901), p. 259, and vol. 3 (1902), p. 383.

I.

	I	II_1	II_2	II_3	II	III	Total
$p>3$ $m>8$	9	$20+p$	$6+2p$	$6+2p$	$32+5p$	23	$64+5p$
$p>3$ $m=8$	8	$20+p$	$6+2p$	$6+2p$	$32+5p$	23	$63+5p$
$p>3$ $m=7$	6	$20+p$	$6+2p$	$6+2p$	$32+5p$	23	$61+5p$
$p=3$ $m>8$	9	23	12	12	47	16	72
$p=3$ $m=8$	8	23	12	12	47	16	71
$p=3$ $m=7$	6	23	12	12	47	16	69

The number of groups of order p^m, with cyclic subgroups of orders p^m, p^{m-1}, and p^{m-2}, are given in Table II:

II.

	p^m	p^{m-1}	p^{m-2}
$p>2$ $m>5$	1	2*	11†

The number of groups of order p^α, $\alpha = 1, 2, 3, 4, 5$ are given in Table III:

III.

	$\alpha=1$	$\alpha=2$‡	$\alpha=3$‡	$\alpha=4$‡	$\alpha=5$§
$p=2$	1	2	5	14	51
$p=3$	1	2	5	15	66
$p=12k-1$	1	2	5	15	$65+2p$
$p=12k+5$	1	2	5	15	$67+2p$
$p=12k-5$	1	2	5	15	$69+2p$
$p=12k+1$	1	2	5	15	$71+2p$

* BURNSIDE, *Theory of Groups*, Art. 65, p. 75.

† Ibid., Art. 66, p. 77; MILLER, Transactions, vol. 2 (1901), p. 259, and vol. 3 (1902), p. 383.

‡ YOUNG, *On the determination of groups whose order is a power of a prime*, American Journal of Mathematics, vol. 15 (1893), pp. 124–178. COLE and GLOVER, *On groups whose orders are the product of three prime factors*, American Journal of Mathematics, vol. 15 (1893), pp. 191–220. HÖLDER, *Die Gruppen der Ordnungen*, p^3, pq^2, pqr, p^4, Mathematische Annalen, vol. 43 (1893), pp. 301–412.

§ BAGNERA, *La composizione dei gruppi finiti il cui grado è la quinta potenza di un numero primo*, Annali di Matematica, vol. 3 (1898), pp. 137–228.

I give in this paper the whole investigation of Class I, but only the results of Classes II, and III. The investigation as a whole will appear in the Publications of the University of Pennsylvania, Mathematics, no. 3.

Class I.

1. *General notations and relations.* — The group G is generated by the two operators P and Q_1. For brevity we set*

$$Q_1^a P^b Q_1^c P^d \ldots = [a, b, c, d, \ldots].$$

Then the operators of G are given each uniquely in the form

$$[y, x] \qquad \begin{pmatrix} y = 0, 1, 2, \ldots, p^2 - 1 \\ x = 0, 1, 2, \ldots, p^{m-2} - 1 \end{pmatrix}.$$

We have the relation

$$Q_1^{p^2} = P^{kp^2}. \tag{1}$$

There is in G, a subgroup H_1 of order p^{m-2}, which contains $\{P\}$ self-conjugately.† The subgroup H_1 is generated by P and some operator $Q_1^y P^x$ of G; it then contains Q_1^y and is therefore generated by P and $Q_1^{p^2}$; it is also self-conjugate in $H_2 = \{Q_1^p, P\}$ of order p^{m-1}, and H_2 is self-conjugate in G.

From these considerations we have the equations

$$Q_1^{-p^2} P\, Q_1^{p^2} = P^{1+kp^{m-4}}, \tag{2}$$‡

$$Q_1^{-p} P\, Q_1^{p} = Q_1^{\beta p^2} P^{\alpha_1}, \tag{3}$$

$$Q_1^{-1} P\, Q_1 = Q_1^{\beta' p} P^{\alpha_1'}. \tag{4}$$

2. *Determination of H_1. Derivation of a formula for $[yp^2, x]^s$.* — From (2), by repeated multiplication we obtain

$$[-p^2, x, p^2] = [0, x(1 + kp^{m-4})];$$

and by a continued use of this equation we have

$$[-yp^2, x, yp^2] = [0, x(1 + kp^{m-4})^y] = [0, x(1 + kyp^{m-4})] \quad (m > 4),$$

and from this last equation,

$$[yp^2, x]^s = \left[syp^2, x\left\{s + k\binom{s}{2} yp^{m-4}\right\}\right]. \tag{5}$$

3. *Determination of H_2. Derivation of a formula for $[yp, x]^s$.* — It follows from (3) and (5) that

---

* With J. W. YOUNG, *On a certain group of isomorphisms*, American Journal of Mathematics, vol. 25 (1903), p. 206.

† BURNSIDE: *Theory of Groups*, art. 54, p. 64.

‡ BURNSIDE: *Theory of Groups*, art. 56, p. 66.

$$[-p^2, 1, p^2] = \left[\beta\frac{\alpha_1^p-1}{\alpha_1-1}p^2,\ \alpha_1^p\left\{1+\frac{\beta k}{2}\frac{\alpha_1^p-1}{\alpha_1-1}p^{m-4}\right\}\right] \quad (m>4).$$

Hence, by (2),

$$\beta\frac{\alpha_1^p-1}{\alpha_1-1}p^2 \equiv 0 \pmod{p^3},$$

$$\alpha_1^p\left\{1+\frac{\beta k}{2}\frac{\alpha_1^p-1}{\alpha_1-1}p^{m-4}\right\}+\beta\frac{\alpha_1^p-1}{\alpha_1-1}hp^2 \equiv 1+kp^{m-4} \pmod{p^{m-3}}.$$

From these congruences, we have for $m>6$

$$\alpha_1^p \equiv 1 \pmod{p^3}, \qquad \alpha_1 \equiv 1 \pmod{p^2},$$

and obtain, by setting

$$\alpha_1 = 1+\alpha_2 p^2,$$

the congruence

$$\frac{(1+\alpha_2 p^2)^p-1}{\alpha_2 p^3}(\alpha_2+h\beta)p^3 \equiv kp^{m-4} \pmod{p^{m-3}};$$

and so

$$(\alpha_2+h\beta)p^3 \equiv 0 \pmod{p^{m-4}},$$

since

$$\frac{(1+\alpha_2 p^2)^p-1}{\alpha_2 p^3} \equiv 1 \pmod{p^2}.$$

From the last congruences

$$(\alpha_2+h\beta)p^3 \equiv kp^{m-4} \pmod{p^{m-3}}. \tag{6}$$

Equation (3) is now replaced by

$$Q_1^{-p}PQ_1^p = Q_1^{\beta p^2}P^{1+\alpha_2 p^2}. \tag{7}$$

From (7), (5), and (6)

$$[-yp, x, yp] = \left[\beta xyp^2,\ x\{1+\alpha_2 yp^2\}+\beta k\binom{x}{2}yp^{m-4}\right].$$

A continued use of this equation gives

$$[yp, x]^s = \left[syp+\beta\binom{s}{2}xyp^2,\right.$$
$$\left. xs+\binom{s}{2}\left\{\alpha_2 xyp^2+\beta k\binom{x}{2}yp^{m-4}\right\}+\beta k\binom{s}{3}x^2yp^{m-4}\right]. \tag{8}$$

4. *Determination of* G. — From (4) and (8),

$$[-p, 1, p] = [Np, a_1^p+Mp^2].$$

From the above equation and (7),

$$a_1^p \equiv 1 \pmod{p^2}, \qquad a_1 \equiv 1 \pmod{p}.$$

Set

$$a_1 = 1+a_2 p,$$

and equation (4) becomes

$$Q_1^{-1} P Q_1 = Q_1^{hp} P^{1+a_2 p}. \tag{9}$$

From (9), (8) and (6),

$$[-p^2, 1, p^2] = \left[\frac{(1+a_2p)^{p^2}-1}{a_2p} bp,\ (1+a_2p)^{p^2}\right],$$

and from (1) and (2),

$$\frac{(1+a_2p)^{p^2}-1}{a_2p} bp \equiv 0 \pmod{p^3},$$

$$(1+a_2p)^{p^2} + bh\frac{(1+a_2p)^{p^2}-1}{a_2p} p \equiv 1 + kp^{m-4} \pmod{p^{m-3}}.$$

By a reduction similar to that used before,

$$(a_2 + bh)p^3 \equiv kp^{m-4} \pmod{p^{m-3}}. \tag{10}$$

The groups in this class are completely defined by (9), (1) and (10).

These defining relations may be presented in simpler form by a suitable choice of the second generator Q_1. From (9), (6), (8) and (10),

$$[1, x]^{p^3} = [p^3, xp^3] = [0, (x+h)p^3] \qquad (m>6),$$

and, if x be so chosen that

$$x + h \equiv 0 \pmod{p^{m-6}},$$

$Q_1 P^x$ is an operator of order p^3 whose p^2 power is not contained in $\{P\}$. Let $Q_1 P^x = Q$. The group G is generated by Q and P, where

$$Q^{p^2} = 1, \qquad P^{p^{m-2}} = 1.$$

Placing $h = 0$ in (6) and (10) we find

$$\alpha_2 p^3 \equiv a_2 p^3 \equiv kp^{m-4} \pmod{p^{m-3}}.$$

Let $\alpha_2 = \alpha p^{m-7}$, and $a_2 = ap^{m-7}$. Equations (7) and (9) are now replaced by

$$\begin{aligned} Q^{-p} P Q^p &= Q^{\beta p^2} P^{1+\alpha p^{m-5}}, \\ Q^{-1} P Q &= Q^{hp} P^{1+ap^{m-6}}. \end{aligned} \tag{11}$$

As a direct result of the foregoing relations, the groups in this class correspond to the partition $(m-3, 3)$. From (11) we find

$$[-y, 1, y] = [byp,\ 1 + ayp^{m-6}] \qquad (m>8).*$$

* For $m=8$ it is necessary to add $a^2\binom{y}{4}p^4$ to the exponents of P and for $m=7$ the terms $a\left(a + abp/2\right)\binom{y}{2}p^3 + a^3\binom{y}{3}p^3$ to the exponent of P, and the term $ab\binom{y}{2}p^2$ to the exponent of Q. The extra term $27ab^2k\binom{y}{3}$ is to be added to the exponent of P for $m=7$ and $p=3$.

It is important to notice that by placing $y = p$ and p^2 in the preceding equation we find that

$$b \equiv \beta \pmod p, \; a \equiv \alpha \equiv k \pmod{p^3} \; (m > 7).^*$$

A combination of the last equation with (8) yields

$$[-y, x, y] = [bxyp + b^2\tbinom{x}{2}yp^2, \tag{12}$$
$$x(1 + ayp^{m-6}) + ab\tbinom{x}{2}yp^{m-5} + ab^2\tbinom{x}{3}yp^{m-4}] \quad (m > 8).\dagger$$

From (12) we get

$$[y, x]^s = [ys + by\{(x + b\tbinom{x}{2}p)\tbinom{s}{2} + x\tbinom{s}{3}p\}p,$$
$$xs + ay\{(x + b\tbinom{x}{2}p + b^2\tbinom{x}{3}p^2)\tbinom{s}{2} \tag{13}$$
$$+ (bx^2p + 2b^2x\tbinom{x}{2}p^2)\tbinom{s}{3} + bx^2\tbinom{s}{4}p^2\}p^{m-6}] \quad (m > 8).\ddagger$$

5. *Transformation of the Groups.*— The general group G of Class I is specified, in accordance with the relations (2) (11) by two integers a, b which (see (11)) are to be taken mod p^3, mod p^2, respectively. Accordingly setting

$$a = a_1p^\lambda, \; b = b_1p^\mu,$$

where

$$dv[a_1, p] = 1, \; dv[b_1, p] = 1 \qquad (\lambda = 0, 1, 2, 3; \; \mu = 0, 1, 2),$$

we have for the group $G = G(a, b) = G(a, b)(P, Q)$ the generational determination:

$$G(a, b): \begin{cases} Q^{-1}PQ = Q^{b_1p^{\mu+1}}P^{1+a_1p^{m+\lambda-6}}, \\ Q^{p^3} = 1, \; P^{p^{m-3}} = 1. \end{cases}$$

---

*For $m = 7$, $ap^2 - a^2p^3/2 \equiv ap^2 \pmod{p^4}$, $ap^3 \equiv kp^3 \pmod{p^4}$. For $m = 7$ and $p = 3$ the first of the above congruences has the extra terms $27(a^3 + ab\beta k)$ on the left side.

†For $m = 8$ it is necessary to add the term $a\tbinom{y}{2}xp^4$ to the exponent of P, and for $m = 7$ the terms $x\{a(a + abp/2)\tbinom{y}{2}p^2 + a^3\tbinom{y}{3}p^3\}$ to the exponent of P, with the extra term $27ab^2k\tbinom{y}{3}x$ for $p = 3$, and the term $ab\tbinom{y}{2}xp^2$ to the exponent of Q.

‡For $m = 8$ it is necessary to add the term $\frac{1}{2}axy\tbinom{s}{2}[\frac{1}{3}y(2s-1) - 1]p^4$ to the exponent of P, and for $m = 7$ the terms

$$x\left\{\frac{a}{2}\left(a + \frac{ab}{2}p\right)\left(\frac{2s-1}{3}y - 1\right)\tbinom{s}{2}yp^2 + \frac{a^3}{3!}\left(\tbinom{s}{2}y^2 - (2s-1)y + 2\right)\tbinom{s}{2}yp^3\right.$$
$$\left. + \frac{a^2bxy^2}{2}\tbinom{s}{3}\frac{3s-1}{2}p^3 + \frac{a^2b}{2}\left(\frac{s(s-1)^2(s-4)}{4!}y - \tbinom{s}{3}\right)yp^3\right\},$$

with the extra terms

$$27abxy\left\{\frac{bk}{3!}[\tbinom{s}{2}y^2 - (2s-1)y + 2]\tbinom{s}{2} + x(b^2k + a^2)(2y^2 + 1)\tbinom{s}{3}\right\}$$

for $p = 3$, to the exponent of P, and the terms

$$\frac{ab}{2}\left\{\frac{2s-1}{3}y - 1\right\}\tbinom{s}{2}xyp^2$$

to the exponent of Q.

Not all of these groups, however, are distinct. Suppose that

$$G(a, b)(P, Q) \sim G(a', b')(P', Q'),$$

by the correspondence

$$C = \begin{bmatrix} Q, & P \\ Q_1', & P_1' \end{bmatrix},$$

where

$$Q_1' = Q'^{y'} P'^{x'p^{m-6}}, \text{ and } P_1' = Q'^{y} P'^{x},$$

with y' and x prime to p.

Since

$$Q^{-1} P Q = Q^{p} P^{1+ap^{m-6}},$$

then

$$Q_1'^{-1} P_1' Q_1' = Q_1'^{bp} P_1'^{+ap^{m-6}},$$

or in terms of Q', and P'

$$[y + b'xy'p + b'^2\binom{x}{2}y'p^2, x(1 + a'y'p^{m-6}) + a'b'\binom{x}{2}y'p^{m-5} + a'b'^2\binom{x}{3}y'p^{m-4}] = [y + by'p, x + (ax + bx'p)p^{m-6}] \qquad (m>8)$$

and

$$by' \equiv b'xy' + b'^2\binom{x}{2}y'p \pmod{p^2}, \tag{14}$$

$$ax + bx'p \equiv a'y'x + a'b'\binom{x}{2}y'p + a'b'^2\binom{x}{3}y'p^2 \pmod{p^3}. \tag{15}$$

The necessary and sufficient condition for the simple isomorphism of these two groups $G(a, b)$ and $G(a', b')$ is, that the above congruences shall be consistent and admit of solution for x, y, x' and y'. The congruences may be written

$$b_1 p^{\mu} \equiv b_1' x p^{\mu'} + b_1'^2\binom{x}{2}p^{2\mu'+1} \pmod{p^2},$$

$$a_1 x p^{\lambda} + b_1 x' p^{\mu+1} \equiv y'\{a_1' x p^{\lambda'} + a_1' b_1'\binom{x}{2}p^{\lambda'+\mu'+1} + a_1' b_1'^2\binom{x}{3}p^{\lambda'+2\mu'+2}\} \pmod{p^3}.$$

Since $dv\,[x, p] = 1$ the first congruence gives $\mu = \mu'$ and x may always be so chosen that $b_1 = 1$.

We may choose y' in the second congruence so that $\lambda = \lambda'$ and $a_1 = 1$ except for the cases $\lambda' \geqq \mu + 1 = \mu' + 1$ when we will so choose x' that $\lambda = 3$.

The type groups of Class I for $m > 8$ * are then given by

$$G(p^{\lambda}, p^{\mu}) : Q^{-1}PQ = Q^{p^{1+\mu}} P^{1+p^{m-6+\lambda}},\ Q^{p^2} = 1,\ P^{p^{m-3}} = 1, \tag{I}$$

$$\begin{pmatrix} \mu = 0, 1, 2\,;\ \lambda = 0, 1, 2, \lambda \leqq \mu\,; \\ \mu = 0, 1, 2\,;\ \lambda = 3. \end{pmatrix}$$

Of the above groups $G(p^{\lambda}, p^{\mu})$ the groups for $\mu = 2$ have the cyclic sub-

* For $m = 8$ the additional term ayp appears on the left side of congruence (14) and $G(1, p^2)$ and $G(1, p)$ becomes simply isomorphic. The extra terms appearing in congruence (15) do not effect the result. For $m = 7$ the additional term ay appears on the left side of (14) and $G(1, 1)$, $G(1, p)$, and $G(1, p^2)$ become simply isomorphic, also $G(p, p)$ and $G(p, p^2)$.

group $\{P\}$ self-conjugate, while the group $G(p^3, p^2)$ is the abelian group of type $(m-3, 3)$.

Class II.

1. *Sections.* — Class II is divided into three sections.
2. *Types.* — The types of Section I ($m > 5$) are defined by

$$(\mathrm{II}_1)\begin{cases} Q^{-1}PQ = Q^{\beta\lambda}P^{1+\alpha p^{m-5}}, \\ R^{-1}PR = Q^{\lambda p}P^{1+\alpha p^{m-4}}, \\ R^{-1}QR = Q^{1+dp}P^{\gamma p^{m-4}}, \\ R^p = Q^{p^2} = P^{p^{m-3}} = 1, \end{cases}$$

where the constants have the values given in the table below,

(II_1)

	a	b	α	β	c	d		a	b	α	β	c	d
1	0	0	1	0	0	1	10	1	0	p	0	0	0
2	0	0	1	0	0	0	11	0	0	p	1	0	0
3	0	0	1	1	0	0	12	0	0	p	0	1	0
4	0	0	1	0	1	0	13	1	0	p	1	0	0
5	0	1	1	0	κ	0	14	0	1	p	0	κ	0
6	0	0	1	1	1	0	15	0	0	p	1	1	0
7	ω	0	p	0	0	1	16	0	0	0	0	0	0
8	0	0	0	0	0	1	17	0	1	0	0	0	0
9	0	0	p	0	0	0							

$\kappa = 0$, 1, and a non-residue (mod p),

$\omega = 0, 1, 2, \cdots, p-1$.

The types of Section 2 are defined by

$$(\mathrm{II}_2)\begin{cases} R^{-1}PR = P^{1+kp^{m-4}}, \\ Q^{-1}PQ = RP^{1+\epsilon p^{m-5}}, \\ Q^{-1}RQ = Q^{p}RP^{\gamma p^{m-4}}, \\ R^p = Q^{p^2} = P^{p^{m-3}} = 1, \end{cases}$$

where the values of the constants are given in the following table:

(II_2)

	k	ε	e	e
1	1	1	0	0
2	0	1	1	0
3	0	1	0	ω
4	0	p	ω	0
5	0	p	1	0
6	0	p	0	κ

$\kappa = 0, 1,$ and a non-residue (mod p),

$\omega = 0, 1, 2, \cdots, p-1.$

The groups of Section 3 are defined by

$$(\mathrm{II}_3)\begin{cases} Q^{-1}PQ = PR^{\gamma 1+\epsilon p^{m-3}}, \\ Q^{-1}RQ = RP^{ep^{m-4}}, \\ R^{-1}PR = Q^p P, \\ R^p = Q^{p^2} = P^{p^{m-3}} = 1, \end{cases}$$

with the values $\gamma = 1$ and a non-residue (mod p); $\epsilon = 1, e = 0, 1, 2, \cdots, p-1,$ and $\epsilon = p$, $e = 0, 1,$ and a non-residue (mod p), $2p + 6$ groups in all.

Class III.

Types. — The types of Class III are defined by

$$(\mathrm{III})\begin{cases} Q^{-1}PQ = P^{1+kp^{m-4}}, \\ R^{-1}PR = Q^\beta P, \\ R^{-1}QR = QP^{\alpha p^{m-4}}, \\ S^{-1}PS = R^\gamma Q^\delta P^{1+\epsilon p^{m-4}}, \\ S^{-1}QS = R^\varepsilon Q, \\ S^{-1}RS = R P^{\beta p^{m-4}}, \end{cases}$$

where the constants have the values given in the table:

(III)

	α	β	c	γ	δ	k	ε	j		α	β	c	γ	δ	k	ε	j
1	0	0	0	0	0	0	0	0	10	0	0	0	0	1	1	0	0
2	1	0	0	0	0	0	0	0	11	1	0	0	0	0	0	1	0
3	κ	1	0	0	0	0	0	0	12	κ	1	0	0	0	0	1	0
4*	0	1	0	1	0	0	0	0	13	0	1	0	0	0	0	0	1
5	0	0	1	0	0	0	0	0	14*	0	1	0	1	0	0	0	1
6	0	0	0	0	0	1	0	0	15	0	0	1	0	0	0	0	1
7*	0	1	0	1	0	1	0	0	16*	0	0	1	0	1	0	0	1
8	0	0	1	0	0	1	0	0	17*	0	1	0	κ	0	1	0	1
9*	0	0	1	0	1	1	0	0									

$\kappa = 0$, 1, and a non-residue (mod p).

* These groups for $p = 3$ ($k \neq 0$) are simply isomorphic with groups in Class II.

UNIVERSITY OF PENNSYLVANIA,
February 23, 1905.

ON THE INVARIANT SUBGROUPS OF PRIME INDEX*

BY

G. A. MILLER

The totality formed by all the operators of any group (G) which are common to all the invariant subgroups of prime index (p) constitutes a characteristic subgroup, and the corresponding quotient group is the abelian group of order p^λ and of type $(1, 1, 1, \cdots)$.† The number of the invariant subgroups of index p is therefore $p^\lambda - 1/p - 1$. The given totality includes all the operators of G which are pth powers, and it is composed of such operators whenever G is abelian. In this case λ is clearly equal to the number of the invariants in the Sylow subgroup of order p^m contained in G. This fact follows also directly from the theory of reciprocal groups, since p^λ is the order of the subgroup generated by all the operators of order p contained in G. For instance, the abelian group G contains only one subgroup of index p whenever its Sylow subgroup of order p^m, $m > 0$, is cyclic, it contains $p + 1$ such subgroups whenever this Sylow subgroup involves two invariants, $p^2 + p + 1$ whenever there are three invariants, etc.

When G is non-abelian the determination of λ is much more difficult. Its value can clearly not exceed the value of λ for a Sylow subgroup of order p^m contained in G, but it may be less. Hence it is of fundamental importance to determine the values of λ for the groups of order p^m, and we shall assume that this is the order of G in what follows. Instead of determining the value of λ for given types of groups, it seems much more desirable to determine the possible types of groups when the value of λ is given. This problem soon becomes very difficult. When $\lambda = 1$, G is cyclic; and when $\lambda = m$, G is the abelian group of type $(1, 1, 1, \cdots)$. These extreme cases relate to fundamental groups whose elementary properties are well known. The case when $\lambda = m - 1$ includes the Hamiltonian groups. All the groups which belong to this case have recently been determined.‡ The main object of the present paper is the study of an interesting category of groups which belong to the case when

* Presented to the Society (San Francisco) February 25, 1905. Received for publication March 2, 1905.

† BAUER, Nouvelles Annales, vol. 19 (1900), p. 509.

‡ Mathematische Annalen, vol. 60 (1905).

$\lambda = m - 2$; viz., all those whose commutator subgroup is of order p. The two abelian groups which come under this case are of types $(3, 1, 1, \cdots)$ and $(2, 2, 1, 1, \cdots)$ respectively. In what follows it will be assumed that G is non-abelian.

The following theorem is of considerable interest in itself and is proved here because it will be convenient to employ a special case of it in the study of the groups under consideration.

THEOREM. *If the commutator subgroup of a group of order p^m is of order p, each of the operators which are common to every subgroup of index p is invariant under the entire group.*

As an invariant subgroup of order p is composed of invariant operators under a group of order p^m, it follows that all the commutators of the group in question are invariant. Hence each operator of the group contains at most p conjugates. If an operator (s) were not commutative with each of the operators which are common to every subgroup of index p, the subgroup of index p composed of all the operators which are commutative with s would not include all the given common operators. As this is impossible each of the common operators is invariant under the entire group.

Let K represent the largest subgroup of G which is common to all its subgroups of index p. Since G contains just $1 + p + p^2 + \cdots + p^{m-3}$ subgroups of this index, the order of K is p^2. We shall consider the groups in question in two sections, — the first will be devoted to those in which K is cyclic and the second to those in which K is non-cyclic. The commutator subgroup of order p will be denoted by C in both of these sections.

§ 1. *Determination of all the possible groups when K is cyclic.*

Let I represent the commutator quotient group of G, that is, I is the quotient group of order p^{m-1} corresponding to C. Since I is abelian and contains just $1 + p + p^2 + \cdots + p^{m-3}$ subgroups of index p it is of type $(2, 1, 1, \cdots)$. The subgroup of G which corresponds to all the operators of I whose orders do not exceed p will be denoted by H. This subgroup is characteristic since it is composed of all the operators of G whose orders are less than p^3. This result follows also directly from the fact that H is composed of all the operators of G whose pth powers are commutators of G.

Since H is a characteristic subgroup of G any two G's must be distinct whenever their H's are distinct. The converse is not necessarily true; that is, two G's involving the same H may be distinct groups as will be seen in what follows. Moreover, H is one of the system of groups mentioned above in which $\lambda = m_1 - 1$; p^{m_1} being the order of H. The value of m_1 is $m - 1$ as stated above. Hence H may be regarded as known and it remains only to find the other possible operators of G. When H is abelian it is simply isomorphic

with I and it is easy to prove that only one group (G_1) can involve such an H and satisfy the conditions that K is cyclic and that there are just $1+p+p^2+\cdots+p^{m-3}$ subgroups of index p, p^m being the order of the group.

The group of cogredient isomorphisms of G_1 is of order p^2 since C is of order p and H is abelian. The truth of this statement may be seen as follows. If t is any operator of $G_1 - H^*$ it has just p conjugates under G_1 and hence it is commutative with a subgroup of order p^{m-2} contained in H. Since these p^{m-2} operators are commutative with every operator of H as well as with t they are commutative with every operator of G_1, that is, the group of cogredient isomorphisms of G_1 is of order p^2. Since the p^{m-2} invariant operators include K they constitute an abelian group of type $(2, 1, 1, \cdots)$.

It is now very easy to prove that there could not be two distinct G's involving the same abelian H. For in two such G's the H's could be made simply isomorphic in such a way that the K of one G would correspond to the K of the other, and that the subgroups composed of the invariant operators of the two H's would correspond. After this had been done, two operators of order p^3, one from each G, could be so chosen that their pth powers would correspond and that they would transform corresponding operators into corresponding operators. That is, the G's would be simply isomorphic.†

When H is non-abelian the considerations become a little more difficult. The results are as follows: *There are always just two G's for a given H except when H involves only p^2 invariant operators; in this special case there is only one G which involves a given H.* The main theorems which will be employed to arrive at these results are: The order of the group of cogredient isomorphisms of G is an even power of p, and the order of the group of cogredient isomorphisms of H is either the same as that of G or it is the quotient obtained by dividing that of G by p^2. This theorem is a direct consequence of the fact that the commutator subgroup of G is of order p, and hence requires no proof here.‡ The other fundamental theorem is as follows: It is possible to make H simply isomorphic with itself in such a way that a subgroup of index p corresponds to any arbitrary subgroup of this index provided these two subgroups involve the same number of invariant operators and include invariant operators of order p^2. Moreover, the subgroups formed by the invariant operators can be made simply isomorphic in any arbitrary manner such that the commutator subgroup corresponds to itself.

The proof of this theorem is readily obtained by means of the following known properties of H. There is one and only one H whose group of cogredient isomorphisms is of order $p^{2\alpha}$, $\alpha < (m-2)/2$; when $2\alpha < m-3$ this H

* This symbol represents the operators of G_1 which are not also in H.

† Bulletin of the American Mathematical Society, vol. 3 (1897), p. 218.

‡ FITE, Transactions of the American Mathematical Society, vol. 3 (1902), p. 342.

is the direct product of a group which has just p^2 invariant operators and an abelian group of type $(1, 1, \cdots)$. If a subgroup of index p in H includes all the invariant operators under H, the order of its group of cogredient isomorphisms is obtained by dividing the order of the group of cogredient isomorphisms of H by p^2. When this condition is not satisfied the subgroup of index p has the same group of cogredient isomorphisms as H has.*

By employing these theorems it is easy to prove that there is only one G which involves H and has the same group of cogredient isomorphisms as H has. For if there were two such G's they could be obtained by adding successively to H the two operators t_1, t_2 of order p^3. As the pth powers of these operators would be invariant under H it would be possible to make H simply isomorphic with itself in such a manner that t_1^p would correspond to t_2^p. Hence the entire groups would be simply isomorphic; that is, there is only one such G.

When the group of cogredient isomorphisms of G is not the same as that of H, it is necessary that H include at least p^3 invariant operators. Whenever this condition is satisfied it is evident that a G can be constructed. If there were two G's involving the same H they could again be constructed by adding to H two operators of order p^3, which may be represented respectively by t_1, t_2. As these operators would be commutative with all the operators of a subgroup of index p contained in H, and as H could be made simply isomorphic with itself in such a manner that these subgroups would correspond to themselves, and, moreover, t_1^p would correspond to t_2^p, it follows as before that there is only one such G. The results of this section may be summarized as follows:

When H is given it is always possible to construct two G's, provided H contains more than p^2 invariant operators. When H involves only p^2 invariant operators, only one G is possible. One of these two systems is composed of all the groups of order p^m which involve invariant operators of order p^3 and contain just $1 + p + p^2 + \cdots + p^{m-3}$ subgroups of index p. *There are just* $\frac{1}{2}(m-1)$ *such groups when* m *is odd. When* m *is even their number is* $\frac{1}{2}(m-2)$. The other system includes the same number of groups when m is even, but it contains one less when m is odd. That is, *the number of the groups which belong to this system is* $\frac{1}{2}(m-3)$ or $\frac{1}{2}(m-2)$, *as* m *is odd or even.* This system is composed of all the possible groups of order p^m which contain just $1 + p + p^2 + \cdots + p^{m-3}$ subgroups of index p and are generated by operators of order p^3 without also including invariant operators of this order.

If a group belongs to the latter system and contains more than p^2 invariant operators it is the direct product of a G which contains just p^2 invariant operators and an abelian group of type $(1, 1, \cdots)$. Similarly, if a group belongs to the former system and contains more than p^3 invariant operators it is the direct product of a G containing just p^3 invariant operators and an abelian group of type $(1, 1, 1, \cdots)$.

*Mathematische Annalen, loc. cit.

§ 2. *Determination of all the possible groups when K is non-cyclic.*

The commutator quotient group I is again the abelian group of type $(2, 1, 1, \cdots)$ and of order p^{m-1}. The characteristic subgroup H corresponds to all the operators of I whose orders do not exceed p just as in the preceding section. All the operators of $G - H$ are of order p^2 while in the preceding section the corresponding operators were of order p^3. As all the operators of H may be of order p it may be the abelian group of type $(1, 1, 1, \cdots)$. If it is not this group it contains just $1 + p + p^2 + \cdots + p^{m-3}$ subgroups of index p and hence it belongs to the known systems of groups which were mentioned above. The pth power of its operators are commutators of G.

Since any two invariant operators of H which are of order p without being commutators can be made to correspond in some simple isomorphism of H with itself it follows that there is one and only one G which has the same group of cogredient isomorphisms as the non-abelian H contained in it. That is, every non-abelian H which includes at least $p^2 - 1$ invariant operators of order p is contained in one and only one G whenever G includes invariant operators not contained in H. As the H's are known and these groups are obtained by extending H by means of an operator of order p^2 whose pth power is in H and which is commutative with every operator of H, this system of groups is completely determined. It therefore remains only to consider those G's whose group of cogredient isomorphisms are larger than those of the H's which they contain.

Every group in question may be constructed by extending some H by means of an operator t which is commutative with each of the operators of a subgroup H' of index p contained in H. As t^p is contained in H' and is commutative with every operator of G without being a commutator of G it follows that H' includes at least $p^2 - 1$ invariant operators of order p. Since the order of the group of cogredient isomorphisms of G is larger than that of H the subgroup H' cannot include all the invariant operators of H. Hence such an H must involve at least p^3 invariant operators under H, and whenever an H has this property a G can evidently be constructed. Since two G's must be distinct whenever their H's are distinct and we found the conditions which H must satisfy in order that a G can be constructed, it remains only to determine how many G's involve the same H. That is, if H is given it is required to find all the possible operators of order p^2 which may be used to extend H in such a way as to give rise to distinct groups.

Let t_1, t_2 be two operators which may be used to extend a fixed H so as to give rise to two distinct G's. Since it is possible to make H simply isomorphic with itself in such a way that any invariant operator of order p which is not a commutator can be made to correspond to any other operator having these properties, it is clear that the two G's would be simply isomorphic if the H' of the

one could be made to correspond to the H' of the other in some simple isomorphism of H with itself. That is, our problem is reduced to finding the number of subgroups of H which could be used as H' but could not be made simply isomorphic in any simple isomorphism of H with itself. In other words, the construction of all the possible groups in question is reduced to the determination of properties of a known system of groups.

As there are just three non-abelian H's which have the same group of cogredient isomorphisms * our problem is reduced to the determination of all the G's which involve these three distinct H's. Two of these contain no invariant operators of order p^2 while the third contains invariant operators of this order. It is not difficult to see that there is only one G which involves a given H that does not include invariant operators of order p^2. For the H' of such a G cannot include invariant operators of order p^2 since it cannot involve all the invariant operators of H. Hence all the subgroups which could be used for H' can be made to correspond in some simple isomorphism of H with itself. That is, there is only one G which involves such an H. It now remains only to consider the case when H includes invariant operators of order p^2 and has a given group of cogredient isomorphisms.

When such an H has more than p^3 invariant operators it follows directly from the known properties of H that it includes three distinct subgroups of index p which can be used as H' since they have the same group of cogredient isomorphisms as H has. These three H''s are simply isomorphic with the three possible H's which have the same group of cogredient isomorphisms. When the H under consideration has only p^3 invariant operators there are only two such G's. Hence we have finally that *there are five G's which have the same group of cogredient isomorphisms provided the order of this group is less than p^{m-3}. When this order is p^{m-3} there are only four such G's.*

As these results follow directly from the known properties of H and from a known theorem in regard to simply isomorphic groups,† it seems unnecessary to enter upon further details. The special case when H is abelian presents no difficulty as there is only one possible G when H is of type $(1, 1, 1, \cdots)$, and there are two G's when H is of type $(2, 1, 1, \cdots)$. The infinite systems of groups determined in this paper are of interest on account of their elementary properties and also in view of the fact that they are completely defined by the number of their subgroups of index p and the order of the commutator subgroup. The close contact between this paper and the one in which the properties of every possible H are determined should be noted. The latter will shortly appear in the Mathematische Annalen under the title: *An Extension of the Hamiltonian Groups.*

* Mathematische Annalen, loc. cit.

† Transactions of the American Mathematical Society, vol. 3 (1902), p. 342.

ON A GENERAL METHOD FOR TREATING TRANSMITTED MOTIONS AND ITS APPLICATION TO INDIRECT PERTURBATIONS*

BY

ERNEST WILLIAM BROWN

The mathematical treatment of any physical problem demands the construction of an ideal problem in which the conditions are different from those of the actual problem. It is assumed that the same general laws hold for the actual and ideal problems, but the complexity of the circumstances surrounding the former makes simplifications of some kind necessary in order that the analysis should not be unreasonably tedious. The ideal problem is therefore usually constructed by neglecting at first some of the influences which form a part of it but which are assumed to affect the results to a much smaller extent than those we retain. The simplified problem, which I call problem A, is then solved. The second step consists in finding what changes are necessary in the solution of problem A when some or all of the neglected influences are included; this second problem I call B. The question under consideration here is the deduction of the solution of B when that of A has been found.

One of the methods for solving B is that known as the *Variation of Arbitrary Constants*. In the deduction of the solution of A, arbitrary constants a arise, and numerical values, depending on the given conditions of the problem, can be assigned to them. In order to obtain the solution of B, it is assumed that the solution of A is still available provided that a different set of values be assigned to the constants a. But the new set of values will usually be different at different times and will therefore depend on the time; in other words, the constants a become variables in problem B.

This last fact gives us another view of the relation between A and B. Suppose that the position of the system which constitutes B is defined by variables x, and that we desire to change from the variables x to other variables p in order to simplify the treatment of B. Then one p-set, namely, the a-set, can be obtained by supposing that the x-set and the p-set are connected by the same equations which connect the x-set and the a-set in the solution of problem A.

* Presented to the Society under a different title, February 25, 1905. Received for publication April 8, 1905.

Thus A is no longer looked upon as a dynamical problem but as a collection of relations which give certain transformations useful for the treatment of B.

But we can go further than this. There is no special necessity imposed on us to use any of the a-sets thus found. We may be able to discover another set of variables, say a ρ'-set, allied to the x-set by relations different from those which connect the x-set and the a-set — relations suggested indeed by the latter and perhaps possessing similar properties, but possessing also additional properties more useful for the purpose in hand.

It is proposed to investigate here one such ρ'-set. The relations between the x- and a-sets usually contain a certain number of other constants α which were present in the original equations of problem A. It is assumed that the x- and ρ'-sets are connected by relations of the same *form* as those which connect the x- and a-sets, but that the α have given *variable* instead of given *constant* values assigned to them in the ρ'-set. Certain advantages are gained in the application of the method to the consideration of the 'indirect inequalities' in the motion of the moon, i. e., in the consideration of those perturbations of the lunar motion which arise from the action of a planet on the earth's motion and which are transmitted to the moon on account of the predominating influence of the earth on the moon. I have given some indications of this ρ'-set in a former paper.* Unfortunately, the main idea was obscured by the method adopted to develop it, and some statements concerning the applicability of it to the lunar problem require further limitations than those previously imposed. The present method of treatment also brings to light a theorem (art. 10 below) which I believe is new and which is of considerable importance in the discussion of the indirect inequalities.

1. *The method of Jacobi.* The equations of a dynamical system are supposed to have been expressed in the canonical form

$$\frac{dx_i}{dt} = \frac{\partial \phi}{\partial y_i}, \qquad \frac{dy_i}{dt} = -\frac{\partial \phi}{\partial x_i} \qquad [\phi = \phi(x_i, y_i, t)], \tag{1}$$

where, to fix ideas, the n variables x_i may be thought as the coördinates defining the position of the system, the n variables y_i as the corresponding momenta, and ϕ as the sum of the kinetic and potential energies; these interpretations are nevertheless not necessary.

Let the variables† x, y be changed to $2n$ new variables p, q, connected with x, y by the relations

* *On the variation of the arbitrary and given constants in dynamical equations*, Transactions of the American Mathematical Society, vol. 4 (1903), pp. 333–350.

† The suffixes will be omitted whenever no confusion is thereby caused. In all cases i, j take the values $1, 2, \cdots, n$.

$$y_i = \frac{\partial S}{\partial x_i}, \qquad q_i = \frac{\partial S}{\partial p_i} \qquad [S = S(x, p, t)], \tag{2}$$

then JACOBI has shown that the equations satisfied by the new variables are

$$\frac{dq_i}{dt} = \frac{\partial}{\partial p_i}\left(\phi + \frac{\partial S}{\partial t}\right), \qquad \frac{dp_i}{dt} = -\frac{\partial}{\partial q_i}\left(\phi + \frac{\partial S}{\partial t}\right), \tag{3}$$

in which, after $\partial S/\partial t$ has been formed from S expressed in its original form as a function of t and the x, p, the function $\phi + \partial S/\partial t$ is supposed to be expressed in terms of p, q, t by means of (2). No assumption is made regarding the form of S, but in order to secure the independence of the x, y and also that of the p, q, it is necessary that the determinant

$$\left| \frac{\partial^2 S}{\partial x_i \partial p_j} \right| \qquad (i, j = 1, 2, \cdots, n),$$

shall not be zero.

Suppose that the x, y receive two sets of independent arbitrary variations δx, δy and $\delta' x$, $\delta' y$, inducing corresponding variations in the p, q. Equations (2) give

$$\Sigma_i(\delta x_i \delta' y_i - \delta y_i \delta' x_i) = \Sigma_i(\delta q_i \delta' p_i - \delta p_i \delta' q_i).$$

Whence, if the x, y are expressed in terms of t and the p, q, and if we adopt the usual notation:

$$[\alpha, \beta] = \Sigma_i\left[\frac{dy_i}{d\alpha}\frac{dx_i}{d\beta} - \frac{dx_i}{d\alpha}\frac{dy_i}{d\beta}\right], \tag{4}$$

we obtain the following relations:

$$\begin{aligned} &[p_i, q_i] = 1, \qquad [p_i, q_j] = 0 \qquad (i \neq j), \\ &[p_i, p_j] = 0, \qquad [q_i, q_j] = 0 \qquad (i = \text{or} \neq j). \end{aligned} \tag{5}$$

We also have

$$\frac{d\phi}{dt} = \frac{\partial}{\partial t}\phi(x, y, t).$$

2. *Particular forms of S.* Suppose first that S is so chosen as to satisfy the equation $\phi + \partial S/\partial t = 0$. The p, q are then constants and they may be considered as the arbitraries of the solution. Such forms of S are solutions, containing n arbitrary constants exclusive of that additive to S, of the partial differential equation

$$\phi\left(x, \frac{\partial S}{\partial x}, t\right) + \frac{\partial S}{\partial t} = 0.$$

Next suppose that S is so chosen as to satisfy the equation

$$\phi + \frac{\partial S}{\partial t} = B, \tag{6}$$

where B is a function of the p only; this is a form much more convenient in many of the applications. Then the p and therefore the q are constants. Let us put

$$p_i = c_i, \qquad q_i = w_i = b_i t + \epsilon_i, \qquad b_i = -\frac{\partial B}{\partial c_i}, \tag{7}$$

the arbitrary constants being c, ϵ, and B being a function of the c only; at present B is undetermined but it may be determined so as to satisfy future conditions as to the forms of x, y, when the latter are expressed in terms of c, ϵ, t.

I shall suppose that this problem has been solved and that this second form for S has been used so that the solution is expressed in terms of c_i, w_i, t, all of which are independent. On account of (2), we have,

$$y_i = \frac{\partial S}{\partial x_i}, \qquad w_i = \frac{\partial S}{\partial c_i} \qquad [S = S(x, c, t)], \tag{8}$$

$$\begin{aligned} [c_i, w_i] &= 1, \quad [c_i, w_j] = 0 && (i \neq j), \\ [c_i, c_j] &= 0, \quad [w_i, w_j] = 0 && (i = \text{or} \neq j). \end{aligned} \tag{9}$$

3. *The constants present in* ϕ. Suppose that ϕ contains m constants α_h to which numerical values have not been assigned. Some or all of these constants will be also present in every form of S. I propose to investigate in this article certain properties of S and ϕ with relation to these constants. It is not necessary to suppose that the α_h are all the constants present in ϕ. Such of them as are not specified in the α_h can be supposed to have had numerical values assigned to them, either explicit or implicit, so that they do not enter into the discussion.

In order to make the following results more general, I suppose that the m constants α_h have been replaced by m variables or constants u_h connected with the α_h by the m independent relations

$$f_k(u_h, \alpha_h, t) = 0 \qquad (h, k = 1, 2, \cdots, m), \tag{10}$$

so that the u_h are functions of the α_h and t only.

The function S was originally a function of the x, c, α, t and it now becomes a function of the x, c, u, t; denote it by S' when it is expressed in the latter form. We have

$$\frac{\partial S}{\partial x_i} = \frac{\partial S'}{\partial x_i}, \qquad \frac{\partial S}{\partial c_i} = \frac{\partial S'}{\partial c_i}, \qquad \frac{\partial S}{\partial t} = \Sigma_h \frac{\partial S'}{\partial u_h} \dot{u}_h + \frac{\partial S'}{\partial t}.$$

* The values 1, 2, $\cdots$, m are given to h, k in all cases.

Put

$$U_h = \frac{\partial S'}{\partial u_h}.$$

Then

$$\frac{dU_h}{dt} = \Sigma_i \frac{\partial^2 S'}{\partial u_h \partial x_i} \dot{x}_i + \Sigma_k \frac{\partial^2 S'}{\partial u_h \partial u_k} \dot{u}_k + \frac{\partial^2 S'}{\partial u_h \partial t}. \tag{11}$$

Also denoting by Q' the expression of any function Q in terms of the x, c u and t, and putting

$$\frac{\partial' \phi}{\partial u_h} = \frac{\partial}{\partial u_h} \phi(x, y, u, t),$$

we obtain

$$\begin{aligned}
\frac{\partial' \phi}{\partial u_h} &= \frac{\partial \phi'}{\partial u_h} - \Sigma_i \frac{\partial \phi'}{\partial y_i} \frac{\partial y_i'}{\partial u_h} = -\frac{\partial}{\partial u_h}\left(\frac{\partial S}{\partial t} + B\right)' - \Sigma_i \dot{x}_i \frac{\partial^2 S'}{\partial u_h \partial x_i} \\
&= -\frac{\partial}{\partial u_h}\left(\Sigma_k \frac{\partial S'}{\partial u_k} \dot{u}_k + \frac{\partial S'}{\partial t} + B\right) - \Sigma_i \dot{x}_i \frac{\partial^2 S'}{\partial u_h \partial x_i} \\
&= -\Sigma_k \left(\frac{\partial^2 S'}{\partial u_h \partial u_k} \dot{u}_k + U_k \frac{\partial \dot{u}_k}{\partial u_h}\right) - \frac{\partial^2 S'}{\partial u_h \partial t} - \frac{\partial B}{\partial u_h} - \Sigma_i \frac{\partial^2 S'}{\partial u_h \partial x_i} \dot{x}_i.
\end{aligned}$$

Combining with (11) we find

$$\frac{dU_h}{dt} = -\frac{\partial' \phi}{\partial u_h} - \Sigma_k U_k \frac{\partial \dot{u}_k}{\partial u_h} - \frac{\partial B}{\partial u_h}, \tag{12}$$

an equation which enables us to find U_h directly.

Again

$$U_h = \frac{\partial S'}{\partial u_h} = \frac{dS}{du_h} - \Sigma_j y_j \frac{dx_j}{du_h},$$

where S and the x are expressed in terms of the c, w, u, t. Hence if U_h be similarly expressed,

$$\frac{dU_h}{dc_i} = \frac{d^2 S}{dc_i du_h} - \frac{d}{dc_i} \Sigma_j y_j \frac{dx_j}{du_h}. \tag{13}$$

But since the c, u, w are independent,

$$\begin{aligned}
\frac{d^2 S}{dc_i du_h} &= \frac{d^2 S}{du_h dc_i} = \frac{d}{du_h}\left(\frac{\partial S}{\partial c_i} + \Sigma_j \frac{\partial S}{\partial x_j} \frac{dx_j}{dc_i}\right) \\
&= \frac{d}{du_h}\left(w_i + \Sigma_j y_j \frac{dx_j}{dc_i}\right) = \frac{d}{du_h} \Sigma_j y_j \frac{dx_j}{dc_i}.
\end{aligned}$$

Substituting in (13), performing the differentiations with respect to u_h, and remembering the notation (4), we find

$$\frac{dU_h}{dc_i} = [u_h, c_i]. \tag{14}$$

In an exactly similar manner it can be shown that

$$\frac{dU_h}{dw_i} = [u_h, w_i], \tag{15}$$

$$\frac{dU_h}{du_k} - \frac{dU_k}{du_h} = [u_h, u_k].$$

The U_h can therefore be determined directly from the expressions for x, y when the solution has been completed.

4. *Two limitations.* In the remaining portion of this paper, I suppose

1°. That t is not present explicitly in $\phi(x, y, u, t)$, so that if t was present explicitly in the original expression for ϕ, one or more of the u must be variable.

2°. That t is not present explicitly in the x, y when the latter have been expressed in terms of the c, w, u, and t.

These assumptions do not very much limit the scope of the results and they render the exposition rather more simple.

We have now

$$\frac{\partial S'}{\partial t} = 0, \qquad \frac{\partial S}{\partial t} = \Sigma_h U_h \dot{u}_h, \qquad \phi + \Sigma_h U_h \dot{u}_h = -B. \tag{16}$$

5. *Variation of arbitrary constants.* The usual method consists in supposing that, after the equations (1) have been solved in terms of the c, w, and t, a new function $R(x, y, t)$ is added to ϕ; account is taken of R by supposing that the x, y retain the same *forms* when expressed in terms of the c, w, and t, so that the c, ϵ are no longer constants. This amounts to nothing more than a transformation from the old variables x, y to the new variables c, w, and the results of art. 1 can therefore be used.

In fact we now have

$$\Sigma_j \left(\frac{dx_i}{dc_j} \dot{c}_j + \frac{dx_i}{dw_j} \dot{w}_j + \frac{\partial x}{\partial t} \right) = \frac{\partial}{\partial y_i} (\phi + R),$$

$$\Sigma_j \left(\frac{dy_i}{dc_j} \dot{c}_j + \frac{dy_i}{dw_j} \dot{w}_j + \frac{\partial y}{\partial t} \right) = -\frac{\partial}{\partial x_i} (\phi + R),$$

which, on account of (9), give

$$\frac{dc_j}{dt} = \frac{\partial Q}{\partial w_j}, \qquad \frac{dw_j}{dt} = -\frac{\partial Q}{\partial c_i},$$

where $Q = \phi + R + \partial S/\partial t$ is supposed to be expressed by means of (8) in terms of the c, w, and t. But ϕ being a function of the x, y, and t, and $\partial S/\partial t$ a function of the x, c, and t, and neither ϕ nor $\partial S/\partial t$ containing total derivatives with respect to t, $\phi + \partial S/\partial t$ must retain the same form whatever be the values given to the c, w. Hence we still have $\phi + \partial S/\partial t = -B$, a function of the c only, and

$$Q = -B + R.$$

Suppose that R is due to changing one of the constants α present in ϕ to a variable. Then

$$Q = -B + \frac{\partial'\phi}{\partial\alpha}\delta\alpha + \tfrac{1}{2}\frac{\partial'^2\phi}{\partial\alpha^2}(\delta\alpha)^2 + \cdots,$$

where $\delta\alpha$ is the variable part added to α. The same method may be pursued, it is true, but if we follow it we take no account of the fact that the *form* of ϕ considered as a function of the x, y, α, t is unchanged, and it would seem as if this unchanged form should be made use of to simplify, if possible, the results. The new method of varying the constants which follows recognizes this fact and brings to light certain results which will be of value in the applications. As in art. 3, the α will be replaced by the u; the two conditions of art. 4 will also be imposed.

6. *Variation of arbitrary and given constants.* I suppose that the x, y retain the same *forms* when the u in ϕ are replaced by functions of the time different from those with which the original problem is solved, so that we replace not only the c, w *but also the* u by their new values in order to get the new values of x, y. The new values of c, w will, therefore, be different in this method from those found by the method of art. 5. In fact we now have, since t is present in the x, y only through its presence in the c, w, u,

$$\sum_j\left(\frac{dx_i}{dc_j}\dot{c}_j + \frac{dx_i}{dw_j}\dot{w}_j\right) + \sum_h \frac{dx_i}{du_h}\dot{u}_h = \frac{\partial\phi}{\partial y_i},$$

$$\sum_j\left(\frac{dy_i}{dc_j}\dot{c}_j + \frac{dy_i}{dw_j}\dot{w}_j\right) + \sum_h \frac{dy_i}{du_h}\dot{u}_h = -\frac{\partial\phi}{\partial x_i}.$$

On account of (9), (14), (15), these give

$$(17)\qquad \begin{aligned} \frac{dw_j}{dt} &= \frac{\partial\phi}{\partial c_j} - \sum_h [c_j, u_h]\dot{u}_h = \frac{\partial}{\partial c_j}(\phi + \textstyle\sum_h U_h\dot{u}_h) = \frac{\partial Q}{\partial c_j}, \\ \frac{dc_j}{dt} &= -\frac{\partial\phi}{dw_j} + \sum_h [w_j, u_h]\dot{u}_h = -\frac{\partial}{\partial w_j}(\phi + \textstyle\sum_h U_h\dot{u}_h) = -\frac{\partial Q}{\partial w_j}, \end{aligned}$$

if u_h be independent of c_j, w_j, and

$$Q = \phi + \Sigma_h U_h \dot{u}_h.$$

The equations are therefore still canonical.

Now when the u, c, w have their original values, denoted by the index 0,

$$\phi^0 + \Sigma_h U_h^0 \dot{u}_h^0 = -B^0. \tag{18}$$

This equation is an identity after the x, y, $\dot{u}$, U have been replaced by their values in terms of the w, c, u, and t, and we can, therefore, give to the c, w, u any arbitrary variations, symbolically denoted by δ_1, so that

$$\phi^0 + \delta_1\phi + \Sigma_h (U_h^0 + \delta_1 U_h)(\dot{u}_h^0 + \delta_1 \dot{u}_h) = -B^0 - \delta_1 B. \tag{19}$$

The actual variations in Q are produced by adding portions, denoted symbolically by δ, to the c, w, u, so that

$$Q = \phi^0 + \delta\phi + \Sigma_h (U_h + \delta U_h) \frac{d}{dt}(u_h^0 + \delta u_h). \tag{20}$$

As ϕ, U, B, $\dot{u}$ are functions of the c, w, u only and not of their derivatives with respect to t, we can take $\delta_1 = \delta$. But we cannot in general put $\delta_1 \dot{u} = d\delta u/dt$. For the identity (19) demands that $\dot{u}$ be first expressed in terms of u by means of (10) and the variation then taken, while in Q a portion δu is to be added to u and the derivative with respect to t then formed. It was the failure to recognize this distinction that caused the error in my former paper.* The method there used implies the substitution of $-B$ for Q under the false assumption that the derivative of δu with respect to t would be the same as $\delta \dot{u}$. The error is immediately evident in the case of those u_h which were originally constants; for them, $\dot{u}_h = 0$ and, therefore, $\delta \dot{u}_h = 0$, while obviously $d\delta u/dt$ will only be zero if δu be constant, and this will not generally be the case.

Combining (19), (20), with $\delta_1 = \delta$, we obtain

$$Q = -B + \Sigma_h U_h \left(\frac{d}{dt} \delta u_h - \delta \dot{u}_h \right), \tag{21}$$

the final form of Q.

I have also obtained this result, when squares and higher powers of δ are neglected, by direct transformation from the ordinary canonical equations of art. 5 above, making use of the relations given in art. 3.

7. There is, of course, no special necessity to confine this method to those variations which are produced by varying the given constants in ϕ. Suppose

* Loc. cit., art. 17.

that the present method is retained for the consideration of a function R added to ϕ, while the given constants in ϕ do not vary. We must then add to Q the function

$$-\delta'\phi + R.$$

the first term of which is necessary to counterbalance the variations of the u in ϕ. But this proceeding appears to be of little or no value, as the results are more complicated than in the usual method. On the other hand the variations of the u are now entirely arbitrary and we might be able to so choose them that the form of Q is more simple or the canonical equations more easy to solve. What advantage may be gained will depend on the particular problem in hand and I shall not pursue this development further in the present paper.

8. *Case of the u_h satisfying canonical equations.* Suppose that the u are divided into two classes c'_h, $w'_h = b'_h t + \epsilon'_h$, where $b'_h = -\partial B'/\partial c'_h$. I suppose that in the initial problem, the c', ϵ' are constants and that B' is a function of the c' only. Instead of using the symbols U, put

$$C'_h = \frac{\partial S'}{\partial c'_h}, \qquad W'_h = \frac{\partial S'}{\partial w'_h} \qquad [S' = S(x_i, c_i, c'_i, w'_i)],$$

so that

$$-\phi - B = \frac{\partial S}{\partial t} = \Sigma_h b'_h W'_h,$$

$$Q = -B + \Sigma_h C'_h \frac{d}{dt}\delta c'_h + \Sigma_h W'_h \left(\frac{d}{dt}\delta w'_h - \delta b'_h\right). \tag{22}$$

Further, I suppose that the variations of the c', w' are given by

$$\frac{dc'_h}{dt} = \frac{\partial}{\partial w'_h}(R' + B'), \qquad \frac{dw'_h}{dt} = -\frac{\partial}{\partial c'_h}(R' + B'),$$

or

$$\frac{d}{dt}\delta c'_h = \frac{\partial R'}{\partial w'_h}, \qquad \frac{d}{dt}\delta w'_h = -\frac{\partial R'}{\partial c'_h} + \delta b'_h.$$

Substituting in (22), we obtain

$$Q = -B + \Sigma_h \left(C'_h \frac{\partial R'}{\partial w'} - W'_h \frac{\partial R'}{\partial c'_h}\right), \tag{23}$$

where it is supposed, in accordance with the previous work, that R' is independent of the c, w.

9. *Properties of the solution in the case of the lunar theory.* The application of these results to the lunar theory is immediate when the necessary limita-

tions which constitute this particular case are imposed. For the sake of brevity, I shall not prove these limitations here but merely state those of them that are necessary for my purpose. They are as follows:

1°. The function R' is the disturbing function due to a planet on the motion of the earth. It can be expanded as a sum of periodic terms and it has a small factor whose square can and will be neglected.

2°. A periodic term in R' of period α years will give rise to periodic terms in the $\delta c'$ of the same period with a factor α in the coefficients: in the $\delta w'$ the factor α^2 also occurs.

3°. It is supposed that the orbit of the earth lies in a fixed plane. There are then two angles w_h', and three angles w_i.

4°. The W_h' consist only of sums of cosines of multiples of the four differences of the five angles w_h', w_i, the terms independent of t being zero; the C_h' consist of sines with similar arguments.

5°. The c_h' are functions of the mean motion and of the eccentricity of the earth's orbit.

10. *The indirect planetary inequalities in the moon's motion.* There are very few sensible planetary inequalities in the moon's motion except when α is a number fairly large compared with unity; only the cases where α is large are treated here. It is further supposed that squares of the small factor which accompanies R' are neglected. Then from (17), (23),

$$\frac{d}{dt}\delta c_i = \frac{\partial Q_1}{\partial w_i}, \qquad \frac{d}{dt}\delta w_i = -\frac{\partial Q_1}{\partial c_i} + \Sigma_j \frac{db_i}{dc_j}\delta c_j + \Sigma_h \frac{db_i}{dc_h'}\delta c_h',$$

where

$$Q_1 = \Sigma_h \left(C_h' \frac{\partial R'}{\partial w_h'} - W_h' \frac{\partial R'}{\partial c_h'} \right).$$

If R' contains a term of long period, α years, the factor α may be introduced in the δc, $\delta c'$ and therefore the factor α^2 in the δw, $\delta w'$, and no higher power of α can be present. Hence the general

THEOREM. *When squares and higher powers of the ratio of the mass of a planet to that of the sun are neglected, the large factor α due to a long period inequality can never occur in the corresponding term in the moon's motion to a power higher than α^2, even if its square is present in the corresponding inequality of the earth's motion.*

The importance of this result lies in the fact that it gives us a means of rejecting in advance many terms of long period, which under the old method would have to be examined in order to find out whether their coefficients were sensible in the moon's motion. The advantage of this method lies, to a great

extent, in the theorem just stated and in the calculation of particular classes of inequalities; this will appear in the following article.

11. *Special classes of inequalities.* I divide the long-period inequalities into three classes for which certain results can be immediately deduced. The *order* of these refers to the power of α present. In the periodic inequalities it is supposed that at least one of the arguments belonging to the disturbing planet is present.

(*a*) *Inequalities containing the lunar arguments.* They are of two kinds. First, those produced by terms of moderately long period in the moon-earth-sun system combined with terms of nearly the same period in the earth-sun-planet system. The former terms have all quite small coefficients and there will be few of such inequalities with sensible coefficients. Second, those produced by combining terms of short period in the two systems. We shall generally be able to neglect the $\delta c'$, $\delta w'$, and the principal parts of the coefficients arise only in the δw_i obtained from limiting the equations to

$$\frac{d}{dt}\delta c_i = \frac{\partial Q_1}{\partial w_i}, \qquad \frac{d}{dt}\delta w_i = \Sigma_j \frac{db_i}{dc_j}\delta c_j.$$

(*b*) *Periodic inequalities not containing the lunar arguments.* Here

$$\delta c_i = 0$$

and the C', W' are limited to those terms which have the arguments w'_h only. The latter have all small coefficients and moreover the terms produced by Q_1 have only the factor α. Hence it will be possible to obtain the chief parts of the coefficients by solving the equations

$$\frac{d}{dt}\delta w_i = \Sigma_h \frac{db_i}{dc'_h}\delta c'_h, \qquad c_i = \text{const.}$$

(*c*) *Non-periodic inequalities.* These again are of two classes. First, the terms containing t to the first power only. These merely give constant additions to the mean motions of the lunar arguments. Second, the secular inequalities, which contain t^2. These latter are determined by

$$\delta c_i = 0, \qquad \frac{d}{dt}\delta w_i = \Sigma_h \frac{db_i}{dc'_h}\delta c'_h,$$

where $\delta c'_h$ is determined from the non-periodic term of R'. This is NEWCOMB's theorem.* The theorem therefore only applies exactly, with the given limita-

* *Action of the Planets on the Moon*, Washington Astronomical Papers, vol. 5 (1894), p. 191.

tions, to the secular inequalities, but it is approximately true for the greater part of the inequalities independent of the lunar arguments, while it is not applicable to inequalities containing the lunar arguments. In arts. 17–19 of my former paper (loc. cit. ante), it was stated that the theorem applied to all ndirect inequalities.

I have contented myself here with a brief outline of the chief results which the method furnishes for the lunar problem. There are numerous details and properties connected with the functions C', W' which might find a place. But these are of more interest in connection with a numerical calculation of the inequalities than of theoretical interest in the variation of the constants or even in the problem of three bodies. It therefore appears advisable to give them later with the publication of the details of a method for such an investigation.

HAVERFORD COLLEGE,
April 8, 1905.

ON HYPERCOMPLEX NUMBER SYSTEMS*

BY

L. E. DICKSON

Generalization of the concept number system.

1. The usual theory relates to systems of numbers $\sum_{i=1}^{n} a_i e_i$ in which the coördinates a_i range independently over all real numbers or else over all ordinary complex numbers; for example, the real quaternion system, or the complex quaternion system. As an obvious generalization,† the coördinates may range independently over all the marks of any field F; for example, the rational quaternion system.

As a further generalization, the sets of coördinates $a_1, \cdots, a_n$ in the various numbers of a system may include only a part of the sets $b_1, \cdots, b_n$, each b_i ranging independently over F; for example, the integral quaternion system. The various coördinates $a_1, \cdots, a_n$ need not have the same range; for example, the numbers

$$(a + 2b\sqrt{2})e_1 + (c + 4d\sqrt{2})e_2 \quad (a, b, c, d \text{ arbitrary integers})$$

form a closed system under addition, subtraction, and multiplication, subject to the associative law, if we set

$$e_1^2 = 2\sqrt{2}\,e_1 - 2e_2, \qquad e_1e_2 = e_2e_1 = e_1, \qquad e_2^2 = e_2.$$

If we make these generalizations on the coördinates, but retain the usual conception of the units e_i, we obtain only subsystems of the usual number systems, the case of modular fields being an exception. It is otherwise if we generalize our conception of the units themselves, freeing them from the restriction‡ of linear independence with respect to the set of all ordinary complex numbers, and assuming merely their linear independence with respect to the given field F.

* Presented to the Society (Chicago), April 22, 1905. Received for publication February 6, 1905.

† DICKSON, *Definitions of a linear associative algebra by independent postulates*, Transactions, vol. 4 (1903), pp. 21–26.

‡ Retained implicitly by TABER, Transactions, vol. 5 (1904), p. 509.

By this generalization we may regard an algebraic field * as a number system; in particular, all ordinary complex numbers form a number system * with respect to the set of all real numbers, the units being $1, i$.

The question of further generalization is considered in § 5.

Closed system of n-tuple elements with respect to a field F.

2. A set of n ordered marks $a_1, \cdots, a_n$ of F will be called an *n-tuple element* a. The symbol $a = (a_1, \cdots, a_n)$ employed is purely positional, without functional connotation. Its definition implies that $a = b$ if and only if $a_1 = b_1, \cdots, a_n = b_n$.

A system of n-tuple elements a in connection with n^3 fixed marks γ_{ijk} of F will be called a *closed system* if the following five postulates hold:

Postulate I. If a and b are any two elements of the system, then $s = (a_1 + b_1, \cdots, a_n + b_n)$ is an element of the system.

Definition. Addition of elements is defined by $a \oplus b = s$.

Postulate II. The element $0 = (0, \cdots, 0)$ occurs in the system.

Postulate III. If 0 occurs, then to any element a of the system corresponds an element a' of the system such that $a \oplus a' = 0$.

THEOREM. *The system is a commutative group under* $\oplus$.

Postulate IV. If a and b are any two elements of the system, then $p = (p_1, \cdots, p_n)$ is an element of the system, where

$$p_i = \sum_{j,k}^{1,\cdots,n} a_j b_k \gamma_{jki} \qquad (i = 1, \cdots, n).$$

Definition. Multiplication of elements is defined by $a \otimes b = p$.

Postulate V. The fixed marks γ satisfy the relations

$$\sum_{j=1}^{n} \gamma_{rtj}\gamma_{jki} = \sum_{j=1}^{n} \gamma_{tkj}\gamma_{rji} \qquad (r, t, k, i = 1, \cdots, n).$$

THEOREM. *Multiplication is associative and distributive.*

To make the system n *dimensional*, we add a sixth postulate:

Postulate VI. If $\tau_1, \cdots, \tau_n$ are marks of F such that $\tau_1 a_1 + \cdots + \tau_n a_n = 0$ for every element $(a_1, \cdots, a_n)$ of the system, then $\tau_1 = 0, \cdots, \tau_n = 0$.

THEOREM. *The system contains n elements* $\epsilon_i = (a_{i1}, \cdots, a_{in})$, $i = 1, \cdots, n$, *such that* $|a_{ij}| \neq 0$.

Indeed, by postulate VI there occurs an element $\neq 0$. Hence if the theorem is false, we may assume that the system contains ν elements $\epsilon_1, \cdots, \epsilon_\nu$, $1 \leqq \nu < n$, such that not every determinant of order ν in the matrix

* For these, PEIRCE's theorem that all but one of the units of the first group can be assumed nilpotent evidently fails. My conception of number systems is therefore wider than TABER's.

$$M_\nu \equiv \begin{bmatrix} a_{11} \cdots a_{1n} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{\nu 1} \cdots a_{\nu n} \end{bmatrix}$$

vanishes, while all of order $\nu+1$ in $M_{\nu+1}$ vanish, $(a_{\nu+11}, \cdots, a_{\nu+1n})$ being an arbitrary element of the system. To fix the notations, let

$$D = |a_{ij}| \neq 0 \qquad (i, j = 1, \cdots, \nu).$$

The expansion of $|a_{ij}| = 0$ $(i, j = 1, \cdots, \nu+1)$ gives a relation

$$\sum_{i=1}^{\nu+1} d_i a_{\nu+1i} = 0, \qquad d_{\nu+1} = D \neq 0,$$

in contradiction with postulate VI.

Every n dimensional closed system a complex number system.

3. This identification may be made formally by establishing a one-to-one correspondence between a and $\sum a_i e_i$, where the e's have the multiplication table $e_j e_k = \sum \gamma_{jki} e_i$.

The following method seems preferable. We define the product* ρa of a mark ρ of F and an element $a = (a_1, \cdots, a_n)$ to be $(\rho a_1, \cdots, \rho a_n)$. In terms of the ϵ_i (end of § 2) any element a is expressible uniquely:

$$a \equiv \sum_{i=1}^{n} \rho_i \epsilon_i, \qquad a_j = \sum_{i=1}^{n} \rho_i a_{ij} \qquad (j = 1, \cdots, n),$$

the ρ's belonging to F. We introduce new units e_i obtained from the ϵ_i by applying the inverse of the transformation (a_{ij}):

$$\epsilon_i \equiv \sum_{j=1}^{n} a_{ij} e_j \qquad (i = 1, \cdots, n),$$

so that $e_1 = (1, 0, \cdots, 0)$; $\cdots$; $e_n = (0, 0, \cdots, 0, 1)$. Then

$$a \equiv \sum_{j=1}^{n} a_j e_j.$$

We note that the e_i are not in general elements of the system; likewise for $(\rho_1, \cdots, \rho_n)$.

On the independence of the postulates.

4. If F has a modulus q, postulates II and III are derivable from the others; in postulate III we may take $a_i' = (q-1)a_i$. If F is the field of integers

* The associative law $\sigma(\rho a) = (\sigma\rho)a$, and the distributive law $\rho(a \oplus b) = \rho a \oplus \rho b$ then follow from these laws for marks and from I. We would give the same definition for $a\rho$ if needed.

modulo q, also postulate IV is redundant, since postulates I and VI insure the presence in the system of every element a, each a_i an integer (§ 2). In fact, there exist integers ρ_i such that

$$\sum_{i=1}^{n} \rho_i a_{ij} \equiv a_j \pmod{q} \qquad (j = 1, \cdots, n).$$

If $n = 1$, $q = 2$ or 3, postulate I is likewise redundant. If $n = 1$, postulate V is redundant. Aside from these special cases, postulates I–VI are independent, as shown by the following systems:

(I) Elements $0, (\pm 1, 0, \cdots, 0), \cdots, (0, \cdots, 0, \pm 1)$; each $\gamma_{ijk} = 0$.

(II) F non-modular; elements a, with each a_i an arbitrary positive integer; each $\gamma_{ijk} = 1$.

(III) Set (II) with element 0 added; each $\gamma_{ijk} = 0$.

(IV) $F \neq GF[q]$; elements a, with each a_i an arbitrary integral mark: each $\gamma_{ijk} = 0$ except γ_{111}, while $\gamma_{111} \neq$ integer.*

(V) Elements a with each a_i arbitrary in F; $n > 1$: $\gamma_{112} = \gamma_{221} = 1$, the remaining γ_{ijk} all zero. [Postulate V fails for $r = t = i = 1$, $k = 2$.]

(VI) Elements a with $a_2 = \cdots = a_n = 0$, a_1 arbitrary; $n > 1$; each $\gamma_{ijk} = 0$. An example for $n \geqq 1$ is the system with the single element 0; each $\gamma = 0$.

Further generalization of the concept number system.

5. The simplest example of a number system the coördinates of whose elements do not all belong to a field is furnished by a reducible system, the coördinates of the numbers of one subsystem belonging to a field of modulus p, those of a second subsystem to a field either without modulus or with modulus $\neq p$. To extend the definition (§ 2) to include this case, we take as elements $(a_1, \cdots, a_n)$, where $a_1, \cdots, a_{n_1}$ belong to a field F_1; $a_{n_1+1}, \cdots, a_{n_1+n_2}$ to F_2; etc. A similar change is to be made in postulate VI. Further, a γ_{jki} vanishes unless j, k, i belong to a single set of subscripts $1, \cdots, n_1$; etc.

Inversely, if in all the numbers $a_1 e_1 + \cdots + a_n e_n$ of a given system, $a_1, \cdots, a_{n_1}$ are marks of a field F_1 of modulus p; $a_{n_1+1}, \cdots, a_{n_1+n_2}$ are marks of a field F_2, $\cdots$, where each F_i $(i > 1)$ is either without modulus or with modulus $\neq p$, then the given system is reducible, the units of the subsystems being $e_1, \cdots, e_{n_1}$; $e_{n_1+1}, \cdots, e_{n_1+n_2}$; etc. Proof will be given here only for the special case† in which each of the fields F_i $(i > 1)$ has a modulus q_i. Let $1 \leqq j \leqq n_1$, $n_1 + 1 \leqq k \leqq n_1 + n_2$. Then

$$p(e_j e_k) = (pe_j)e_k = 0, \qquad q_2(e_j e_k) = e_j(e_k q_2) = 0.$$

* For $n = 2$ we may use the elements $A_1 + A_2 j$; $A_r = a_r + \alpha_r i$, $i^2 = j^2 = -1$, $ji = -ij$; viz. the real quaternion system expressed in two units $1, j$, but without γ's satisfying IV.

† Noted by Mr. WEDDERBURN, who gave a different proof.

Since p and q_2 are distinct primes, there exist integers P and Q such that $Pp + Qq_2 = 1$. Hence $e_j e_k = 0$.

It remains to examine the restriction that the coördinates of the numbers of a system (or of a subsystem of a reducible system) shall belong to a field. One would surely retain the assumption that the aggregate of these coördinates can be extended by adjunction to an aggregate A forming a commutative group under an operation called addition and such that a second operation called multiplication, obeying the associative and distributive laws, can be performed within A. We retain the commutative law for multiplication and assume that a product vanishes only when one factor vanishes, as otherwise in either case the aggregate of coördinates would be fully as complicated as a number system. Under these assumptions the aggregate forms part of a field. *

THE UNIVERSITY OF CHICAGO,
February 5, 1905.

* See KÖNIG, *Algebraischen Grösszen*, 1903, pp. 8–9.

A THEOREM ON FINITE ALGEBRAS*

BY

J. H. MACLAGAN-WEDDERBURN†

FROBENIUS ‡ and C. S. PEIRCE § have shown that, in the domain of all real numbers, the only linear associative algebras every number of which, except zero, possesses an inverse, are quaternions and its subalgebras, and also that in the complex domain no algebra has that property. In the present paper it is shown that the Galois field is the only algebra of this type which possesses a finite number of elements.

§ 1.

Since addition is commutative in a linear associative algebra, it may be shown, as in the Galois field theory,‖ that for any number x of the algebra there exists a prime integer p for which p times x is zero and further that p is the same for every x. It follows that, in the group formed by the numbers of the algebra under addition, every element is of period p and therefore the order of the group is of the form p^n where n is some positive integer. The numbers $\neq 0$ of the algebra thus form a group F, of order $p^n - 1$, under the operation of multiplication. The self-conjugate elements of F, together with the zero element, form a Galois field. For, if y_1 and y_2 are self-conjugate elements of F and x is any element,

$$(y_1 + y_2)x = y_1x + y_2x = xy_1 + xy_2 = x(y_1 + y_2),$$

i. e., the self-conjugate elements are closed under addition, as well as under multiplication, and hence with 0 form a Galois field. Since the identity is a self-conjugate element of F this Galois field always exists. If the order of the Galois field is p^m, the order of the subgroup of F, composed of its self-conjugate elements, is of order $p^m - 1$. The subgroup will be denoted by G and the corresponding Galois field by $GF[p^m]$.

Let x_1 be any element $\neq 0$ of the algebra; then there are exactly p^m distinct

* Presented to the Society (Chicago) April 22, 1905. Received for publication April 5, 1905.

† Carnegie Research Scholar (Scotland).

‡ G. FROBENIUS, Crelle, Bd. 84 (1878), p. 59.

§ C. S. PEIRCE, American Journal of Mathematics, vol. 4 (1881), p. 225.

‖ As developed, for instance, in Professor DICKSON'S *Linear Groups*, p. 9.

numbers of the form $\xi_1 x_1$, where ξ_1 is a mark of $GF[p^m]$. If x_2 is any element not included in this set, there are p^{2m} distinct numbers of the form $\xi_1 x_1 + \xi_2 x_2$. Similarily if x_3 is any number, which is not of the form $\xi_1 x_1 + \xi_2 x_2$, there are p^{3m} numbers of the form $\xi_1 x_1 + \xi_2 x_2 + \xi_3 x_3$, and so on.

We can evidently enumerate in this way all the numbers of the algebra and hence we can find say s numbers $x_\alpha (\alpha = 1, 2, \cdots, s)$ such that any number x can be expressed uniquely in the form

$$x = \sum_{\alpha=1}^{s} \xi_\alpha x_\alpha$$

where $\xi_\alpha (\alpha = 1, 2, \cdots, s)$ are marks of $GF[p^m]$. The order of F is then $p^{ms} - 1$.

§ 2.

Let x_a be any number of the algebra which does not lie in $GF[p^m]$. Then, if y_1 and y_2 are any two numbers commutative with x_a, $y_1 + y_2$ and $y_1 y_2$ are also commutative with x_a and hence the set of all numbers commutative with x_a forms a subalgebra. The group, formed by the numbers of this algebra under multiplication, will be denoted by F_{x_a}. Since F_{x_a} contains G, we find as in § 1 that its order is of the form $p^{ms_a} - 1$, where s_a is some positive integer. Hence on dividing the elements of F into conjugate classes we get

$$p^n - 1 = p^m - 1 + \sum_{\alpha=1}^{t} k_\alpha \frac{p^n - 1}{p^{ms_\alpha} - 1}. \tag{1}$$

This shows that, if the least common multiple of the $s_\alpha (\alpha = 1, 2, \cdots, t)$ is s', $p^m - 1$ is divisible by $(p^n - 1)/(p^{ms'} - 1)$. Therefore

$$(p^m - 1)(p^{ms'} - 1) = l(p^n - 1).$$

Reducing this modulo p^m we see that l must have the form $kp^m - 1$ $(k > 0)$. Since $ms' \leqq n$, we have $k = 1$ and $ms' = n = ms$.

§ 3.

It follows from the theory of hypercomplex numbers, that there is an equation of lowest degree,

$$f(x) \equiv x^r + a_1 x^{r-1} + a_2 x^{r-2} + \cdots + a_{r-1} = 0, \tag{2}$$

with coefficients in $GF[p^m]$, which is satisfied identically by any given number x of the algebra, irrespective of any special relation between the coördinates of x, except the condition that they lie in $GF[p^m]$. Further, there is at least one element of the algebra which satisfies no similar equation of lower degree. Indeed, (2) states that $x^{r-1}, x^{r-2}, \cdots, x^0$ are linearly independent with respect to

$GF[p^m]$, and the condition of independence can evidently be put in a form which states that certain determinants, whose elements are rational integral functions of the coördinates of x, do not all vanish identically. Hence there must be some set of values of the coördinates for which $x^{r-1}, x^{r-2}, \cdots, x^0$ are independent and hence the particular x so obtained can satisfy no equation of lower degree than r. (2) is called the characteristic or identical equation, while the equation of lowest degree satisfied by a given x is called its reduced* equation.

$f(x)$ is irreducible in $GF[p^m]$; for, were it reducible neither factor would possess an inverse, contrary to our hypotheses. Similarily, the reduced equation of a given number is irreducible.

For any given x all the roots of its characteristic equation $f(x) = 0$ are roots of its reduced equation† $\phi(x) = 0$, both being regarded as ordinary equations in the $GF[p^m]$, and therefore, since $\phi(x)$ is irreducible, $f(x)$ is a power of $\phi(x)$ and the degree of the reduced equation is a divisor of r.

We now assume that F_{x_a} is abelian for every x_a which does not lie in the $GF[p^m]$. It is shown below that the general case can be made to depend on this simpler one. Under this assumption F_{x_a} is the multiplicative group of a Galois field and is therefore cyclic.‡ If x_a is chosen as the generator of F_{x_a} and if s_a is the degree of the reduced equation of x_a, there are exactly $p^{ms_a} - 1$ different rational functions of x with coefficients in $GF[p^m]$; hence the order of F_{x_a} is $p^{ms_a} - 1$. Now it was shown above that, for some x_a, $s_a = r$ and also that each s_a is a factor of r. Hence r is the least common multiple of $s_1, s_2, \cdots$ and therefore $r = s$. But $ms = n$. Hence F_{x_a} is identical with F, i. e., the algebra is a Galois field.

Suppose now that F is not abelian. Then for some x, not contained in G, F_x is not abelian. Similarly there must be some element of F_x, which is not self-conjugate in F_x and is such that the group of those elements of F_x which are commutative with it, is not abelian; and so on. We could then deduce in this way a series of groups of decreasing order, no member of the series being abelian. This is however impossible since each group of necessity contains G and the order of F is finite. Hence F must be abelian.

§ 4.

The same result can be deduced as follows without the aid of the theory of the characteristic equation. It was shown in § 2 that, if F is not a field, its elements can be arranged in subgroups the orders of whose multiplicative groups

* FROBENIUS, loc. cit.

† Cf. FROBENIUS, loc. cit.; E. WEYR, Monatshefte für Mathematik und Physik, vol. 1 (1890), p. 163.

‡ Cf. DICKSON, loc. cit., p. 13.

are of the form $p^{ms_a} - 1$. It follows then from SYLOW's theorem, that every prime divisor of $p^n - 1$ is also a divisor of some $p^{ms_a} - 1$ $(ms_a < n)$. This is however only possible in the following two cases:* (i) $p = 2$, $n = 6$; (ii) $p = 2^k - 1$, $n = 2$.

In case (i) we have $m = 1$ since the only divisors of n are 2 and 3, which are relatively prime, hence from (1), § 2,

$$2^6 - 1 = 1 + x_1 \frac{2^6 - 1}{2^3 - 1} + x_2 \frac{2^6 - 1}{2^2 - 1}$$

where x_1 and x_2 are integers not both zero. This gives $62 = 9x_1 + 21x_2$, an equation which cannot be satisfied by integers. Hence this case cannot occur. Case (ii) is evidently inadmissable since $n = 2$ is a prime.

§ 5.

The following proof is perhaps more direct. If all the elements of the algebra are multiplied successively by any one element except zero it is easily seen from the distributive law that they are permuted among themselves in such a way as to leave their additive relations unchanged, i. e., each such operation gives an isomorphism of the additive group with itself. † It follows that in the group of isomorphisms of the additive group there are two subgroups simply isomorphic with the multiplicative group; namely, one obtained by left and the other by right handed multiplication.

Each operation of one of these subgroups is commutative with every operation of the other, and it is easily seen that their greatest common subgroup corresponds to the set of the self-conjugate elements of F and is therefore of order $p^m - 1$. These two subgroups then generate a subgroup of order $(p^n - 1)^2/(p^m - 1)$ of the group of isomorphisms. Since the additive group is an abelian group of type $(1, 1, \cdots, 1)$, its group of isomorphisms is the general linear homogeneous group $GLH(n, p)$ of order $(p^n - 1)(p^n - p)\cdots(p^n - p^{n-1})$, an expression which is divisible by $(p^n - 1)^2/(p^m - 1)$ only in the two special cases mentioned above, unless $m = n$. Having excluded these two cases above, we must have $m = n$. Therefore F is abelian. ‡

THE UNIVERSITY OF CHICAGO,
March 31, 1905.

* BIRKHOFF and VANDIVER, *On the Integral Divisors of* $a^n - b^n$, Annals of Mathematics, vol. 5 (1904), p. 177; L. E. DICKSON, *On the Cyclotomic Function*, American Mathematical Monthly, April, 1905.

† I am indebted to Professor E. H. MOORE for calling my attention to this point of view.

‡ Professor DICKSON has deduced the same result from the theory of canonical forms.

THE RELATION OF THE PRINCIPLES OF LOGIC TO THE FOUNDATIONS OF GEOMETRY*

BY

JOSIAH ROYCE

INTRODUCTION.

In the year 1886, in the Philosophical Transactions of the Royal Society, Mr. A. B. KEMPE published *A Memoir on the Theory of Mathematical Form*, in which, amongst other matters, he discussed the fundamental conceptions both of symbolic logic and of geometry. The ideas there indicated were further developed, by Mr. KEMPE, in an extended paper *On the Relation between the Logical Theory of Classes and the Geometrical Theory of Points*, in the Proceedings of the London Mathematical Society for 1890. Despite the close attention that has since then been devoted to the study of the foundations of geometry, Mr. KEMPE's views have remained almost unnoticed. They concern, however, certain matters which recent research enables us to regard with increasing interest. I have been led, therefore, to attempt a restatement of KEMPE's logical-geometrical theory. The restatement has led me to conceptions which, although implied in those which Mr. KEMPE emphasizes, present a number of aspects which I believe to be novel, so that a considerable part of the present research follows a path of its own. My introductory words will indicate the nature of KEMPE's contribution to the problem of the foundations of geometry, the kind of task which his work has set before me, and my own main interest in preparing this paper.

The fundamental ordinal relation of geometry is the relation which can be, at pleasure, described as the triadic relation "between," or as an asymmetrical, transitive dyadic relation, such as "before," or "antecedent to," or "sequent to." Essentially the same relation is at the root of all serial order, and on this basis the logic of such order has lately been elaborately discussed by Mr. BERTRAND RUSSELL, in his *Principles of Mathematics*.

The axioms of geometry, as Dr. VEBLEN has stated them (Transactions of the American Mathematical Society, July, 1904), consist (1) of

* Presented to the Society April 29, 1905. Received for publication May 5, 1905.

assertions characterizing the "between" relation, and duly restricting the application of this relation so far as the "lines" of geometry are concerned, and (2) of existential propositions defining certain entities that shall possess the relation. A similar prominence of asymmetrical transitive relations appears in Dr. HUNTINGTON'S various *Sets of Postulates* for numbers, groups, etc. (Ibid., January and April, 1905).

The algebra of logic may be viewed (as Dr. HUNTINGTON, following Mr. PEIRCE and SCHROEDER, has lately afresh shown in detail), as depending upon the relation of inclusion or subsumption, sometimes symbolized by $\prec$. This relation is dyadic and transitive, and may be either symmetrical or unsymmetrical. Upon the basis of this one relation we can define the various operations of formal logic, such as logical multiplication and addition. If the relation $\prec$ is in a given instance symmetrical, it ensures what is commonly viewed as the "uniqueness" of an entity. That is: $a \prec a$; and if $a \prec b$, while $b \prec a$, than $b = a$ (see Dr. HUNTINGTON'S paper of July, 1904, in these Transactions, for a fuller statement of the various results of these considerations). The relation $a \prec b$, in so far as it obtains between non-equivalent elements, may serve to define linear series: $a \prec b \prec c \prec d$, etc.; where $a \prec c$, and $a \prec d$. In such a series c may obviously be said to lie "between" b and d, and the analogy to the geometrical relation "between" is in so far plain. "Dense," and in fact, continuous linear series of the subsumption type can be conceived after the analogy of point series. But on the other hand, a system of logical classes differs, with respect to linear relations, from a system of points on a line in *two* very notable ways: —

(1) If $a \prec b \prec d$, and if it is also true that $a \prec c \prec d$, any one of the three relations $b \prec c$, $c \prec b$, $c = b$, is indeed *possible;* but, in case of the logical entities, it is also possible that b and c are such that no one of these relations actually holds between these two. Thus, Siberia is included within the Russian Empire, which itself may be viewed as included within the "Eurasian" continent. And Siberia is also included in Asia, which may also be regarded as included within the "Eurasian" continent. These subsumptions are transitive, and in so far linear in their type. Yet the Russian Empire and Asia do not form a pair possessing the relation $\prec$, read in either sense.

(2) If $a \prec b \prec c \prec d$, and if, also, $i \prec b \prec c \prec j$, the relations of i and a, of j and d are similarly left indeterminate. These relations need not be directly expressible in terms of $\prec$ at all. That is, nothing in the logical relations forbids linear series (whether dense, or continuous or not) to have two or more "points," i. e., elements, in common, while any number of the other elements of the series remain entirely distinct. The logical lines, as Mr. KEMPE observes, may intersect any number of times.

For this very reason, however, the system of logical entities may be viewed

simply as much more general and inclusive than the system of the points of space. And thus it becomes possible to regard a given space-form as *a selection from amongst the entities present in a system that exemplifies the logical relation* $\prec$. That is: One may view the points of a space as a select set of logical elements, chosen, for instance, from a given "universe of discourse." This thought, whose possible fruitfulness for the logical development of the foundations of geometry I regard as highly notable, is the essential thought at the basis of Mr. KEMPE's paper of 1890, cited at the outset of this introduction.

The reason why such a thought seems promising is this: The relations amongst logical entities are, in any case, the most fundamental relations that we know. Experience shows us in the outer world those ordinal space-relations which geometry generalizes in the concept of "between." But our own thinking processes show us the meaning of the logical relation $\prec$. The latter relation, then, is more suited to be the basis for a theory of the logic of an exact science, in case we can only so define and restrict its application that our ideal geometrical relations can come to be viewed as special instances of those forms which we can develop by the use of pure logic.

Mr. KEMPE's procedure, in the paper of 1890, is, in bare outline, as follows: He sets out, not by assuming the ordinary algebra of logic, but by defining, through postulates, a purely abstract set of entities called by him the "base-system," and a relation which may be viewed as a generalized "between." The latter is the relation which, in its most general form, is characteristic of what KEMPE himself calls, later in his paper, "flat collections" of any number the elements of the "base-system." But the relation first appears as a *triadic* relation, and is so characterized in the postulates. KEMPE uses the notation: $ac \cdot b$, to mean the assertion: "a, b, and c, form a 'linear triad,' with b between a and c." So far the expressions used resemble those for Dr. VEBLEN's generalized relation "in the order abc." But KEMPE's linear triad has these fundamental properties: (1) "If $ab \cdot c$, and $a = b$, then $c = a = b$." (2) "If $a = b$, then $ac \cdot b$ and $bc \cdot a$, whatever entity of the system c may be."* In other words, KEMPE permits the "between" relation to hold where the related elements are, for all the purposes of the operations of the system, identical; and then he defines the distinctness of elements by means of a restriction of the relations that are permitted to hold in triads of distinct elements. The result is that the "between" relation becomes Dr. VEBLEN's "in the order," whenever the elements are all distinct.

The other properties of the "between" relation which are in question for

* I vary a little the order of Mr. KEMPE's statement of his principles. The relation $=$ is defined by Mr. KEMPE only in a very highly abstract form which I need not here attempt to discuss. Geometrically interpreted, if this relation holds between points, these become identical points.

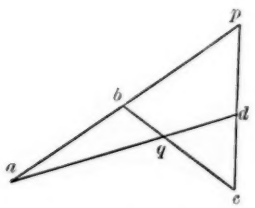
Mr. KEMPE, are obtained by him through assuming two forms of triadic "transversal" propositions as fundamental postulates, viz.: *

I. If two linear triads, $ap \cdot b$ and $cp \cdot d$, exist, such that (as indicated by the notation), b is between a and p, and d is between c and p, then there exists an entity, q, which lies, in a linear triad, between a and d, and, in another linear triad, between b and c.

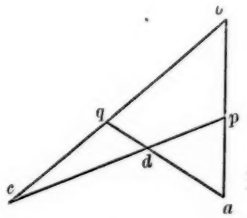
II. If, in the linear triads $ab \cdot p$ and $cp \cdot d$ (as indicated by the notation), p lies between a and b, and d between c and p, then q exists such that q lies, in a linear triad, so that d is between a and q, while, in another linear triad, q is between b and c.

If one interprets these assertions as relating to points in space, they become assertions obviously relating, respectively, to the diagrams following. But, as

I.



II.



they are stated at the outset of Mr. KEMPE's paper, these principles have no specification beyond what the general properties of the linear triad, as just defined, predetermine.

One other existential proposition Mr. KEMPE uses as his *fifth* fundamental principle. This is simply the proposition that any entity belongs to the base-system whose presence there is not inconsistent with the four other principles, —a proposition which of course formally renders the two existential principles, here numbered I and II, superfluous; and which leaves the account of the "base system" inevitably somewhat unsatisfactory.

Mr. KEMPE now proceeds upon this basis, to show, by a decidedly original, although necessarily intricate procedure, that the elements of the base system, as thus defined, possess the properties and relations of a system of logical classes, or of other entities subject to the algebra of logic. In other words, he develops the entire algebra of logic, including the definitions and properties of the operations of logical multiplication and logical addition, without any other assumptions than those simple properties of the "between" relation which have just been stated. The proofs given are such as to apply to any finite number of the elements. Mr. KEMPE leaves, however, some doubt as to infinite collections.

* I vary slightly Mr. Kempe's mode of enunciating these existential propositions at the outset of his paper.

Highly instructive observations are incidental to this development. The system of logical entities appears as possessing a thoroughly symmetrical structure. The "zero"-element and the "universe"-element have no essential distinction from any other similarly related pair of "obverse" elements. Negatives, in general appear as "obverses," because of the symmetrical contrast of their respective relations to the remainder of the system. All the fundamental relations of logic appear as triadic rather than as dyadic. But upon this triadic basis, polyadic relations also develop — the relations of KEMPE's "Flat-collections." These collections, thus named by reason of their resemblance to the various possible configurations of points in an n-dimensional space — "on a line," "on a plane," "in a three dimensional space," etc. — are Mr. KEMPE's means of relating the purely logical to the geometrical entities.

The junction of his principles with the regular algebra of logic once completed, although leaving certain doubts as to the application of his proof to infinite sets, KEMPE proceeds to the geometrical application. By (1) selecting certain "linear sets" of the elements of the base system; and (2) selecting from these sets those which conform to a new principle (here for the first time introduced into the essay), namely, to the principle that any two of the elements of a selected linear set shall determine the whole linear set to be selected, KEMPE is in possession of a system of foundations for a geometry of a "flat space" of n dimensions. The further development of such a geometry is indeed merely sketched in the paper in question. But since the "triangle-transversal" axiom is provided for by the initial principles of the system, and since, by the selection of the linear sets of elements, the ordinary properties of the geometrical "between," and the axiom as to the determination of a line by any two of its points are now also secure, KEMPE's result, although only indicated in his text, is in the main clear. *A space of n dimensions is a select class or set of elements which themselves are entities in a logical "field."* The selection of the entities of a given space is arbitrary; and so the space-forms whose entities are selected may be varied in any way whatever which is consistent with the triangle-transversal-axiom, and with the properties of the generalized between-relation. The problem of the continuity of the geometric sets is only very generally treated, and is not solved.

The wide outlook thus suggested into the theory of space-forms certainly deserves to be better considered than KEMPE's treatment of the subject has so far been considered by mathematicians. For me, however, as a student of philosophy, a still further interest attaches to those results which I have thus suggested, an interest which my mathematical colleagues may also share.

The problem of the foundations of geometry is only a part of that general problem regarding the fundamental concepts of the exact sciences which is now so widely studied. KEMPE's research suggests that, since metrical relations, and

therefore (as KEMPE himself briefly indicated in his *Theory of Mathematical Form*, § 309), the whole algebra of ordinary quantity, can be reduced, in any system of three or more dimensions, to a series of propositions based upon purely ordinal relations, — the entire system of the relationships of the exact sciences stands *in a much closer connection with the simple principles of symbolic logic* than has thus far been generally recognized.

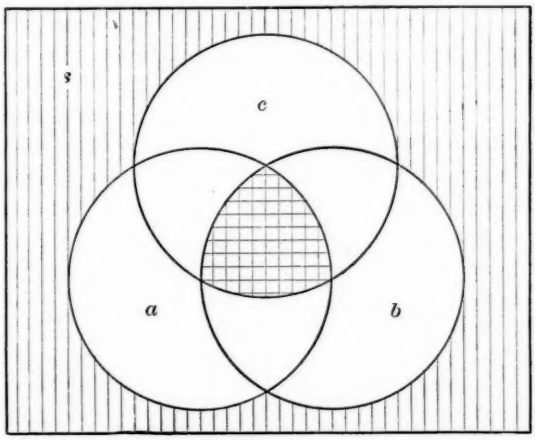
Mr. BERTRAND RUSSELL, using very different methods, insists, indeed, in general, upon the closeness of such a connection. But the distinction between the "logic of relations," and the older "logic of classes," and of "propositions," — a distinction which Mr. BERTRAND RUSSELL in his *Principles of Mathematics* regards as something quite fundamental, seems to me to become, in the light of KEMPE's research, a distinction probably quite superficial. Hence to my mind, Mr. KEMPE's theory goes far deeper than Mr. RUSSELL's. Give us a system of entities of the types of logical classes, and we shall find that their relations (all statable in terms of KEMPE's "between"), are already (quite apart from a separate "logic of relations"), certainly as rich as the totality of the relations studied in geometry, and are, for reasons upon which KEMPE has dwelt, probably as rich as the totality of the relations known to the exact sciences, at least so far as the latter have yet been developed. The bare prospect of such a result deserves a careful consideration, in case one takes interest in the unification of the categories of science. KEMPE's theory promises such an unification.

The present memoir proposes to contribute towards a more precise statement of the theory thus outlined. At the basis of my own discussion, I place, however not KEMPE's "between" relation, but another fundamental relation of symbolic logic which has the interest of being absolutely *symmetrical*, while, when it obtains amongst n entities, it permits (upon the basis of certain simple existential propositions), the definition of the properties of KEMPE's "flat collections," and so the definition of asymmetrical relations of a very high degree of complexity. This change of starting point is the prime novelty of the present discussion, as contrasted with KEMPE's.

The *contrast between symmetrical and unsymmetrical relations* seems, to the ordinary view, absolute. Mr. RUSSELL, in his late volume, so treats it. Geometry, and the ordinary algebra of quantity (as these subjects are usually treated), seem to depend on regarding the distinction as quite fundamental. In symbolic logic, however, as Mrs. LADD-FRANKLIN long ago pointed out (in her paper on the algebra of logic in the volume called "*Studies in Logic by members of the Johns Hopkins University*," Boston, 1883), a "symmetrical copula," namely that of "inconsistency," or of "opposition," can be made to accomplish all the work of the ordinary unsymmetrical copula $\prec$. In other words, if I have otherwise defined the meaning of "not," the statement "x is inconsistent with not-y," means the same as "x implies y." The copula in the former case is

symmetrical, in the latter unsymmetrical. The former expression makes explicit the "relative product" (as it is called by PEIRCE and RUSSELL) of *two* symmetrical relations (viz., "opposes" and "not"). This "relative product" is, itself, indeed an unsymmetrical relation. But the constituents of this product are symmetrical. This already suggests how asymmetry may be definable in terms of symmetry.

Using as my suggestion some brief observations of KEMPE (in §§ 75–82, of his paper in the Proceedings of the London Mathematical Society), I have therefore chosen to define, by postulates, at the outset of my discussion, a symmetrical relation which I may call "the *O*-relation." This relation is essentially polyadic, and applies at once to any number of terms greater than one. In logical terms, this is the relation in which (if we were talking of the possible chances open to one who had to decide upon a course of action) any set of *exhaustive but, in their entirety, inconsistent choices* would



stand to one another. It is also the relation in which the members of any set of areas stand to one another when there is no area (except the "zero"-area) which is common to *all* the areas of the set at once, while together these areas exhaust some larger surface (which therefore resembles, in its relation to them, a logical "universe"). Thus if, in the annexed diagram, the surface s contains three circles, a, b, and c, and if we then agree to disregard, or to view as stricken out or destroyed, the here shaded portions of the diagram, the circles a, b, and c have then only the stricken-out or "zero" area in common, while together they exhaust what we thus permit to remain of the surface s. In consequence, a, b, and c here form what KEMPE calls "an obverse collection," and what I call, in this paper, an "*O*-collection."

If *two* objects stand in the *O*-relation each to the other, then these two are related as a pair of "negative" or contradictory classes, or statements, are related to one another. But any number or multitude of objects, in case such are otherwise permitted to exist in a given system, may stand, not in pairs, but as a whole collection, in this relation, and will then constitute what I call an "*O*-collection." The name that I give to the relation is derived from the close analogy of such collections, even when they contain more than two elements, to KEMPE's pairs of "obverse elements." But I do not myself wish to call the *O*-collections, with KEMPE "obverse collections," because, as will be seen, I find it convenient to use an expression in my text (where I speak of "mutually obverse" collections), in a way that would make such usage confusing. Hence I read the expression "*O*-collection," simply as written.

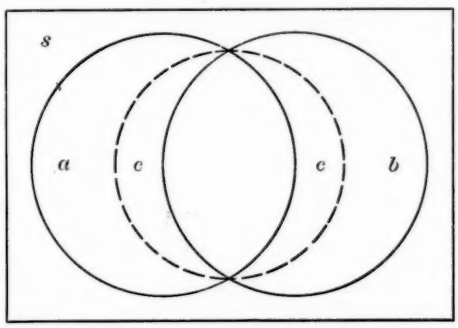
In my text the *O*-relation is entirely freed from dependence upon all such examples as the ones just used, and is defined solely by postulates, and is to be taken solely as it is there defined. The fact however that it is, in its relational properties, identical with the "yes-no" relation — the earliest exact relation defined by the human mind — is, in this introduction, important. *For what I am in the end to show is that all the serial and other ordinal relations known to logic and to geometry, and all the operations known to both, so far as they are pure exact sciences, are ultimately reducible to assertions that certain entities do, while certain entities do not stand to one another in the perfectly symmetrical O-relation.*

My procedure differs from KEMPE's, not only in making this wholly symmetrical relation, instead of KEMPE's "between," fundamental, but also in the existential principles which I assume. KEMPE's "transversal" axioms form with me a theorem, proved late in the discussion. My own existential principles have to be wide enough to provide for the "continuity" of the system, or, rather, for its inclusion of infinitely numerous continuous systems, and definite enough to make the system of the entities to which the logical calculus is applicable a determinate manifold, inclusive of the points of a space of n-dimensions. The usual treatment of the algebra of logic provides only for arbitrarily determined sets of 2^n or of 2^{2^n} entities in a given logical system. KEMPE calls any such selected set a "full set." KEMPE's further postulate, however, calling for "all entities" consistent with the formal laws, is itself indefinite. In seeking adequate postulates I have been led to two observations which, although in themselves fairly obvious, seem to me to be new, viz., (1) a relation is here shown between the existence of logical sums and products and the general theory of limits and of continuity; and (2) a general definition of the pairs of elements which I have called "conjugate resultants"*

* This concept of "conjugate resultants" is generalized from KEMPE's own generalization of his "unsymmetrical resultants" in § 28 of the essay of 1890. My use of the concept differs in many ways from his.

is made centrally important. The algebra of logic, so far as I know, has not hitherto been brought into definite relations with the problem of the continuum. This is one of the things that I here accomplish. This undertaking involves proving all the principles of logic so as to make them applicable to infinite sets of entities at once. This also I have here done.

KEMPE's "linear triad" of elements is represented, in any logical system of classes, by the classes, or areas a, b, and c, which stand in the relation which is represented in the adjoining diagram by the closed figures so lettered. Any area c which includes the common part of a and b, and which is included within their logical sum, is, in KEMPE's phrase, such that "c is between a and b." I hereafter symbolize this relation, in my own way, as $F(c|ab)$. The relation in question is called by me the F-relation, because it is that characteristic of



KEMPE's "flat collections." The F-relation, so long as "obverses" or "negatives" exist, follows immediately from, and is equivalent to, an O-relation. For, in the diagram if s is the total surface in which a, b, and c are included, then when "c is between a and b," "a, b, and $\bar{c}$ ($\bar{c}$ being the obverse of c) constitute an O-collection," or "are in the O-relation."

The outcome of our discussion will show that, while logical relations can be indifferently stated as O-relations, or stated as F-relations, or (when once addition, multiplication, and negation have been defined) can be stated in terms of equivalence, the F-relations are the only natural means of expressing the geometrical ordinal relations. This difference, however, between the logical and the geometrical entities, is due to the simple fact that (as KEMPE points out), when geometrical sets are considered, the obverses of the elements of any set are excluded from that set, so that the obverses may be viewed as ideal elements of the geometry in question. In fundamental meaning all these relations spring from a common root.

If "m is between b and c," I sometimes call m "mediator" or again, on occasion, "resultant" of b and c. I extend the term "resultant" to include

the case where a single element stands in an F-relation to any collection of elements.

The axioms or principles assumed at the outset of my discussion, in § 19, are verifiable for a collection of areas all of which are included within a given area, if the O-relation is interpreted as, for the sake of illustration, I have just done. The consistency of these axioms is thus secured from the start. For the verification of "Principle VI" of my set of principles, see § 118 of my text. For KEMPE's term "base-system," I substitute "the system Σ."

The considerable length of the discussion may be justified by the importance, (1) Of a development of the principles of logic solely in terms of a symmetrical polyadic relation; and (2) Of the need of supplementing KEMPE's results by a theory of the continuity of the "base-system."

THE SYSTEM Σ, AND THE O-COLLECTIONS.

CHAPTER I. DEFINITIONS AND PRINCIPLES.

1. The system Σ, whose structure we are to consider in what follows, consists of certain "elements," which we shall regard, in the present discussion, as simple and homogeneous. As symbols for these elements, we shall employ the small letters of the alphabet: $a, b, \cdots, i, j, \cdots, x, y$, etc. In many cases, for the sake of distinguishing one element of a set from others, we shall need subscript marks; and for these too we shall nearly always employ small letters, or if that be convenient, numbers, thus: $a_k, b_i, \cdots, x_1, x_2, p_r$, etc. It is to be noted however, that, unless the contrary is especially indicated, these subscripts are *merely* convenient distinguishing marks; so that the numbers when used as subscripts will, in general, *not possess any ordinal meaning*, but will be used merely as tags. The few exceptions to this rule will explain themselves.

2. The elements thus symbolized may be viewed either singly, or in their collections. A collection of elements will usually be in question, in what follows, as a sort of complex or secondary unit. We shall apply predicates to collections when they are viewed as such complex units, shall compare collections, combine them into larger collections, make partitions of them into the partial collections of which larger collections are composed, classify collections, etc. A collection may consist of a single element of Σ, and is then called a monad, or a monad-collection. A collection of two elements is called a pair, of three a triad, of four a tetrad, of n elements an n-ad. But a collection may consist of an infinite multitude of elements. And, in fact, whenever our statements, and whenever the conditions imposed in the course of a given investigation, do not set definite limits to the multitude of the elements that belong to such collections as are at any time in question, it is always to be understood that the collections

of which we then speak are actually permitted to consist not only of any finite number, but of any multitude whatever of elements.

3. Collections in which a single element, such as b, or x, is viewed as occurring repeatedly, are to be regarded as permissible objects for our consideration: and if we define the number or the multitude of elements in such a collection, or if, for any other reason, we have to treat all the various elements of a given collection in various ways, then, for the purposes of the count, or of the other individual treatment of the elements of any collection, the various repetitions of a given element are to be treated as, in so far, distinct members of the collection in question. Empty, or "zero" collections will not be considered in the present discussion.

4. In order to symbolize a collection without indicating, by the mere symbol, any assertion except the assertion that the collection consists of certain elements, we shall write the symbols for the elements in question, separated by commas, and in a parenthesis. Thus the symbol (a, b) denotes "the pair which consists of the elements a and b." The symbol (x, x) denotes "the pair consisting of x and of x repeated." The symbol (x, x, y, y, r), denotes "the collection consisting of x and of x repeated, and of y and of y repeated, and of r." The symbol $(x_1, x_2, \cdots, x_r, \cdots)$ denotes "the collection consisting of x_1, of x_2, and of an unspecified multitude of other elements, each of which is symbolized by x written with some subscript." In such a case, if no restriction of the multitude of the elements is stated, this multitude of the elements of the collection need not be limited to that of the whole numbers; and the use of whole numbers as subscripts is then of no special significance, beyond that of the convenience of such subscript-symbols. Other subscript-symbols would be equally possible, and may, upon occasion, be used.

5. In many cases, we shall need to symbolize a collection without at the moment designating any of the single elements of which it consists. In such case we shall use Greek letters, α, β, γ, δ, etc. (and, in a few cases, the capital Greek letters also), as symbols for *entire collections*. Thus, the symbol α means "the collection designated as α, consisting of elements which are not hereby further specified." In such a case, the collection α may be, in fact, a perfectly determinate collection, and the symbol α will then be merely a convenient abbreviation. In other cases, α may stand for an unspecified instance of some class of collections; the members of the collection at any time in question being left, by the conditions of the discussion, to be otherwise determined. A collection α or β or δ may be unrestricted as to the multitude of its elements, or may be a monad, a pair, a triad, etc., according as the conditions of a given statement permit or determine.

6. A frequent operation, in our discussion, will be the adjunction of elements to an already given collection, or to elements already under consideration. The

symbol (α, x) will mean "the collection consisting of the collection α of elements together with the element x, adjoined thereto." The symbol (α, β, γ) will mean "the collection formed by adjoining to the collection α all the elements of the collections β and γ." Greek letters, as symbols for whole collections, and small letters a, x, m, etc., as symbols for elements, may thus be combined in the same expression in order to indicate what adjunctions are at any time in question.

7. When we speak simply of "a collection" of the elements of Σ, without further specification of the character of this collection, *the order in which the elements of the collection are named, or otherwise indicated, or in which they stand in the collection, is wholly indifferent.* A collection, in such a general case, *is determined wholly by the fact that certain elements do, while certain elements do not belong to it.* The arrangement of the elements within the collection will concern us only in case the definition of a given type of collections, or the conditions of a given problem, expressly require us to take note of such arrangement.

8. Fundamental, in our discussion of the properties of the system Σ, is a classification of the collections of elements into those which are, and those which are not what we shall call "O-collections." The O-collections form a class of collections whose fundamental properties we define by the laws hereafter stated. The symbol $O(xyz\cdots)$ is to be read as the statement: "The elements x, y, z, etc., taken together, constitute an O-collection." The symbol $O(\alpha)$ is to be read as the statement: "The collection α is an O-collection." The symbol $O(\beta\gamma)$ is to be read as the assertion that "The total collection formed out of the collections β and γ is an O-collection." At pleasure we shall also use the abbreviated form of expression: "the collection $O(xyz\cdots)$," meaning "the collection such that the assertion $O(xyz\cdots)$ is true." The symbol $O(\alpha x)$ means that "the collection formed by adjoining the element x to the collection α is an O-collection." The symbols $O(\alpha\beta\gamma\delta)$, $O(xy\alpha pq\beta jr\delta)$, etc., are to be read so as to assert that whatever total collection of elements and collections is indicated by the letters enclosed in the parentheses, is an O-collection.

9. If a collection is *not* an O-collection, the fact may have to be separately asserted. We propose in the cases where, for the sake of conciseness such usage is advisable, to call the class of all those collections which are *not* O-collections, E-collections. The symbol $E(\alpha)$ may be read at pleasure as the assertion: "The collection α is *not* an O-collection," or again "is an E-collection." Correspondingly we read the symbols: $E(xy)$ (where we also assert that "x and y form an E-pair," or "do not form an O-pair"); $E(x\alpha)$; $E(xy\beta\gamma)$, etc. If α is not an E-collection, then $O(\alpha)$.

In general, we shall speak of "O-pairs," "O-triads," "O-tetrads," "O-n-ads," and of "E-pairs," "E-triads," etc., wherever our collections, whether E-col-

lections or O-collections, are restricted as to the number of elements in a way to which we wish to call attention.

10. When we simply assert $O(\alpha)$ or $E(\beta)$ of any collection, *the order in which the elements of an O-collection or of an E-collection are named or considered is indifferent.* The elements in question in such cases, simply *do* or *do not* belong to the collections in question without regard to the order in which the elements stand.

Equivalent elements.

11. In case two elements are such that *each of them can be substituted for the other in every O-collection in which that other occurs*, while leaving the collection in question still an O-collection, then these two elements are said to be *equivalent each to the other* or to be *mutually equivalent elements.* But if there exists an O-collection into which either of these elements enters, while the other cannot be substituted for the first, in that collection, without altering it into an E-collection, then the two elements are not equivalent. The equation $x = y$ means, therefore, in the present discussion *not* that x and y are identical, but simply that "either x or y may be substituted for the other in any O-collection wherein that other occurs, while the substitution leaves the collection in question still an O-collection." In case it were possible that *neither* x nor y formed a member of *any* O-collection, this definition would imply that they were then also equivalent. The usual properties of the relation of equivalence obviously follow from this definition: viz., $x = x$; if $x = y$, then $y = x$; if x, y, and z are such that $x = y$, and $y = z$, then $x = z$. It is plain that, if $x = y$, either x or y can be substituted for the other in any E-collection in which that other occurs. For if $x = y$, it is, by definition, impossible that $E(x\alpha)$ while $O(y\alpha)$; hence, if $E(x\alpha)$, $E(y\alpha)$ follows; and the converse is also obvious from the definition of equivalence.

12. As just pointed out, equivalent elements need not be identical. Hence, although the assertion: "a is not equivalent to b," obviously implies the assertion: "a is not the same element as b," — these two assertions must still be carefully distinguished, since the second of them does not imply the first. For the assertion "a is not equivalent to b," we shall use the symbol $a \neq b$. This means that "there exists at least one O-collection into which one of these elements enters, while, if the other is substituted for the first in that collection, the collection in question becomes an E-collection."

Mutually obverse elements and collections.

13. If two elements of Σ, say p and q, are such that $O(pq)$, then p and q are said to be *mutually obverse elements*, or *obverses*, each of the other. Mutually obverse elements are then such elements as together form an O-pair. If the assertion $O(xx)$ were true of any element, that element would be an obverse of itself. A given element may possess various obverses.

14. If a set of O-pairs exists, such that, for certain existent elements of Σ, $O(ab)$, $O(cd)$, $O(ef)$, $O(pq)$, etc. (this set possessing either finite or infinite multitude), and if we suppose a collection α made up by the selection of one member, and of one member only, from each of these pairs, and if a collection β is supposed to be made up out of all the remaining members of the pairs, the collections α and β are said to be mutually *obverse collections*.

15. In consequence of the formation of the collections α and β just defined, certain repetitions of elements may occur in α, or in β, or in both. In such a case, just as in the cases mentioned above, in 3, the various repetitions of any given element are to be regarded as distinct members of the collections α and β in question. Thus, if the O-pair $O(ab)$ is repeated, then, out of the two pairs $O(ab)$ and $O(ab)$, treated, for the purpose in hand, as distinct pairs, we can form, according to the procedure defined in 14, the collections (a, a), and (b, b). These collections are hereupon to be regarded, by virtue of our definition, as mutually obverse collections. Again, if the pairs $O(pq)$, $O(xy)$, $O(mn)$, $O(xk)$, and $O(mq)$, are given, then the collections (p, x, m, x, m), and (q, y, n, k, q), are mutually obverse collections, as are also any two collections that can be formed out of these same pairs by any permissible variation of the procedure defined in 14. The order in which the members of each of the mutually obverse collections are written, is again indifferent.

16. Suppose a collection δ is first given. Let each element of this collection be such that it can be made to form an element of a pair of elements which (whether the elements of this pair are, or are not, repetitions of those present in other pairs), is distinct from the pair of which any other element of δ is a member. If each of the pairs thus formed is an O-pair, then the collection ϵ, consisting of all the remaining elements of the pairs in question, is an obverse of the collection δ, while δ is an obverse of ϵ. Thus, if δ is the collection (x, m, k, l, l), and if $O(xy)$, $O(mj)$, $O(kn)$, $O(lr)$ and $O(ls)$; and if ϵ is the collection (y, j, n, r, s), then the collections δ and ϵ are mutually obverse collections.

Complements and resultants.

17. In case an element q exists such that, for a given collection β, $O(\beta q)$ is true, the element q is called a *complement* of β. In case q and r exist such that, for a given collection β, $O(\beta q)$ is true, while, at the same time $O(qr)$, then r is called a *resultant* of β.

18. The properties of those collections which may be formed of the elements of Σ are, in the main, properties determined by the existence of equivalent and of non-equivalent elements, of obverses and of resultants, together with the existence of certain laws and principles which hold valid for the system.

19. To the statement of these principles or "postulates" we now proceed. They may be classified under two heads. They are: (1) General laws to which

all O-collections, in case such exist, are to conform; and (2) Principles requiring, either conditionally or unconditionally, the existence of certain elements, and of certain collections.

(1) *Laws to which all O-collections conform.*

I. If $O(\alpha)$, then $O(\alpha\gamma)$, whatever collection γ may be.

II. If, whatever element b_u of β be considered, $O(\delta b_u)$, and if $O(\beta)$ is also true, then $O(\delta)$.

(2) *Principles requiring the existence of elements of* Σ.

III. There exists at least one element of Σ.

IV. If an element x of Σ exists, then y exists such that $x \neq y$.

V. Whatever pair (p, q) exists, such that $p \neq q$, r also exists such that, while both $O(rp)$ and $O(rq)$ are false, $O(pqr)$ is true.

VI. If w exists such that $O(\vartheta w)$, then v also exists such that $O(\vartheta v)$, and such, too, that, whatever element t_u of ϑ be considered, $O(vwt_u)$.

20. These principles may be restated, with less use of symbols, thus:

I. An O-collection remains an O-collection, whatever elements or collections may be adjoined to it.

II. If a collection β, consisting wholly of elements which are complements of a collection δ, is an O-collection, then δ itself is an O-collection.

III and IV (in combination). The system Σ contains at least one pair of mutually non-equivalent elements.

V. If any pair of mutually non-equivalent elements is given, a third element of Σ exists which forms an O-pair with *neither* of the elements of this pair, but which is such that the three elements in question together constitute an O-triad.

VI. If there exists any complement of a given collection ϑ, then, if w be such a complement, there exists a complement of ϑ, viz. v, such that every element of ϑ is a complement of the pair (v, w).

At the close of the introduction a system Σ which conforms to all the foregoing principles, has been already pointed out.

Chapter II. Elementary Consequences of the Principles.

21. *The elimination of obverses.* If any collection α is such that x and y exist such that $O(\alpha x)$ and $O(\alpha y)$, while $O(xy)$, then $O(\alpha)$. This follows directly from principle II, if the pair (x, y) be viewed as the collection β of that principle.

22. *The correspondence of mutually obverse O-collections.* If any collection π is such that $O(\pi)$, and if a collection ρ is a collection which is an obverse collection of the collection π, then $O(\rho)$. For let p_u be any element of π. Then in ρ there exists (by the definition of mutually obverse collections, as given

in 14, 15, 16), some element r such that $O(p_\mu r)$. By principle I, we may adjoin to the O-pair, $O(p_\mu r)$, all the remaining elements of ρ besides r, and the thus enlarged collection will still remain an O-collection; so that $O(p_\mu \rho)$. As an analogous result holds of every other element, p_ν, of π, without exception, it appears that π consists entirely of elements each of which is a complement of the collection ρ. Since, however, $O(\pi)$, by principle II, $O(\rho)$.

23. *The elimination of common elements.* If η and ϑ are mutually obverse collections; and if x exists such that $O(\beta x)$ and $O(\vartheta x)$, while y exists such that $O(xy)$, then $O(\beta\eta)$. For, by 22, from $O(\vartheta x)$ follows $O(\eta y)$, since, by adjoining x to ϑ, and y to η, we form the two mutually obverse collections (x, ϑ) and (η, y). By adjunction, in accordance with principle I, from $O(\beta x)$ follows $O(\beta\eta x)$ and from $O(\eta y)$ follows $O(\beta\eta y)$. Since $O(xy)$, there follows, from $O(\beta\eta x)$ and $O(\beta\eta y)$, by principle II, as explained in 21, $O(\beta\eta)$.

24. *The elimination of partial collections.* If ϵ consists solely of elements which are complements of a collection λ, if δ and γ are mutually obverse collections, and if $O(\delta\epsilon)$, then $O(\gamma\lambda)$. For ϵ by hypothesis consists of elements which are complements of λ. Let e be, then, an element of ϵ. Then $O(e\lambda)$, and hence, by adjunction (principle I), $O(e\gamma\lambda)$. Furthermore, δ consists wholly of elements which are complements of γ. Hence if d is an element of δ, $O(d\gamma)$ and hence, by adjunction (principle I), $O(d\gamma\lambda)$. Any element of ϵ, and also any element of δ, is thus a complement of $(\gamma\lambda)$. Hence the O-collection $O(\delta\epsilon)$ consists entirely of elements which are complements of the collection (γ, λ). Hence the collection (γ, λ) is itself an O-collection, by principle II. If ϵ reduces to a single element, e, then the hypotheses above stated reduce to $O(\delta e)$ and $O(\lambda e)$, while δ and γ are mutually obverse collections, and the result then becomes identical with that of 23. But if δ reduces to a single element d and γ to an obverse element c such that $O(cd)$, then the result is that if ϵ consists solely of elements which are complements of a collection λ, and if d is such that c exists such that $O(cd)$, while $O(d\epsilon)$, then $O(c\lambda)$.

25. The operations of the reduction of collections through the elimination of elements and of partial collections as explained in the foregoing, will be found to be of fundamental significance throughout our procedure in what follows.

26. *The existence of obverse elements.* By virtue of principles III and IV, there exists (x, y) such that $x \neq y$. Since $x \neq y$, it follows from the definition of equivalence (11) that there exists at least one O-collection into which one of these elements, say x, enters; while in that collection (whether y is also a member of the collection in question or not), y cannot be substituted for x without changing the collection into an E-collection. Let α be the collection thus characterized. Then, by hypothesis, $O(\alpha)$.

If α contains all the elements of Σ, occurring either once each, or in any multitude of repetitions, then a collection exists which contains all of the ele-

ments of Σ, and which is an O-collection. If α does not contain all of the elements of Σ, nevertheless, by principle I, since $O(\alpha)$, all of the elements of Σ which do not appear in α may be adjoined to α, and the resulting collection, say ϑ, will be an O-collection.

Hence, in any case, there exists a collection which contains every element of Σ (each element occurring in that collection either once only, or else repeatedly), while this collection, say ϑ, is an O-collection.

27. Since ϑ is such that $O(\vartheta)$, every element of Σ is a complement of ϑ, by principle I. Let w be any element of Σ. Since w is a complement of ϑ, it follows by principle VI, that v exists, such that, whatever element x of Σ, or of ϑ, be chosen, $O(xvw)$. Since every element of ϑ is thus a complement of the pair (v, w), while $O(\vartheta)$, it follows by principle II, that $O(vw)$ is true. By adjoining to the pair (v, w) all the elements of Σ which do not belong to this pair, we now have $O(\Sigma)$, an assertion according to which each element of Σ is supposed to appear once, and without repetition, in the O-collection in question. Since $O(vw)$, v and w are mutually obverse elements.

Since any element whatever of Σ may be taken instead of w, while, each time, an element would be found to take the place here occupied by v, we have so far two results:

(1) The system Σ, taken in its entirety, is an O-collection.

(2) Every element of Σ possesses at least one obverse.

28. A fundamental property of all pairs of mutually obverse elements hereupon comes to our notice, and is as follows: Let b be any element. Let q and r be two obverses of b, so that $O(qb)$ and $O(rb)$. Hereupon let γ be any collection such that $O(\gamma q)$; that is, let the collection $O(\gamma q)$ be any O-collection into which q, one of the obverses of b, enters. Then, by adjunction (principle I), we have, since $O(br)$, $O(br\gamma)$; and, since $O(\gamma q)$, $O(qr\gamma)$ (wherein we may of course change, as we here do, at pleasure, the order in which the members of the O-collection are written). The collection (r, γ) is thus such that, if either of the members of the O-pair $O(bq)$ be separately adjoined to it, the resulting enlarged collection is each time an O-collection. Hence, by principle II, $O(r\gamma)$. It thus appears that, whatever the collection γ may be, if $O(q\gamma)$, $O(r\gamma)$ follows. By a precisely analogous reasoning we could show that if γ is such that $O(r\gamma)$, $O(q\gamma)$ follows. Hence the two obverses of b here in question, viz., q and r, are such that either of them may be substituted for the other in any O-collection in which that other occurs, while still leaving that collection an O-collection. Hence by the definition of equivalence $q = r$. As the reasoning thus used applies to any two obverses of the same element b, whatever b is, we have, as a result, the proposition that *any two obverses of the same element are mutually equivalent elements.* That is, again, if q is an obverse of b, and b is an obverse of r, then $q = r$.

29. If $x = y$, while $O(qx)$ and $O(ry)$, then, by the definition of equivalence, $O(rx)$, and from $O(qx)$ and $O(rx)$ follows, by 28, the proposition $q = r$. Hence *all the obverses of equivalent elements are equivalent.*

30. If $u \neq v$, while $O(nu)$ and $O(mv)$, then $m \neq n$. For if $m = n$, then, by 29, $u = v$. Hence *the obverses of non-equivalent elements are themselves non-equivalent.*

31. It thus appears that all the obverses of the same element, or of equivalent elements, are mutually equivalent, and that non-equivalent elements cannot possess mutually equivalent obverses, still less the same obverse. Accordingly, since equivalence here means capacity for mutual substitution in O-collections, we may henceforth let a *single one* of the obverses of a given element represent the whole class of these obverses, for all the purposes involved in the present discussion of O-collections. This uniquely selected representative of all the obverses of any element x, we shall henceforth regard, therefore, as *the* obverse of x, and as equivalent to the obverse of any element equivalent to x. We shall symbolize this single representative of all these obverses by $\bar{x}$, or, in general, by writing a bar above the symbol of the element of which at any time we define the obverse. $\bar{x}$ cannot be equivalent to the obverse of any element which is not equivalent to x. As the unique representative of the obverses of $\bar{x}$ we may hereupon take an element symbolized by $\overline{\bar{x}}$.

32. By definition, and by 31, $x = \overline{\bar{x}}$; and $\overline{\bar{x}}$ will henceforth be so chosen as to be identically the *same* element as x. The operation of obversion (that is, of finding, for any element x, the unique representative, $\bar{x}$, of the class of elements any one of which forms, with x, an O-pair), hereupon becomes an entirely univocal operation. This operation, if once repeated, is so defined as to be an operation which restores to us the original element.

33. When one collection, δ, is an obverse of another collection, ϵ (see 14, 15, 16), each of these collections, by the substitution of the equivalent elements (in case such substitution is required for the purpose), may be made into a collection consisting wholly of the unique representatives of the obverses of the various members of the other collection. An obverse of the collection δ, thus reduced to the form of a collection of the unique representative obverses of the elements of δ, shall henceforth be symbolized, in our discussion, by $\bar{\delta}$. By the symbol $\bar{\delta}$ we mean, therefore, a certain chosen unique representative of all those collections any one of which is an obverse of the collection δ. The collection $\bar{\delta}$ may then be so chosen as to be identical with δ.

34. If $a = b$, then $O(a\bar{b})$ and $O(b\bar{a})$. For $O(a\bar{a})$ by definition (13, 31). Hence, since $a = b$, and since, by the definition of equivalence, we can accordingly substitute b for a in $O(a\bar{a})$, we have $O(b\bar{a})$. And since $O(b\bar{b})$, we have, by the substitution of a for b, $O(a\bar{b})$. On the other hand, if either $O(a\bar{b})$, or $O(b\bar{a})$, is known to be true, then, by the definition of obverses, a is

an obverse of $\bar{b}$, or b is an obverse of $\bar{a}$, as the case may be. But in either case, since obverses of the same element or equivalent elements are equivalent, $\bar{a} = \bar{b}$, and $a = b$.

35. A fundamental characteristic of the system Σ is, further, the fact that: *No monad is an O-collection.* For if $O(x)$, then, by adjunction (principle I), $O(xx)$. By the same principle, if $O(x)$, $O(xq)$, whatever element q may be. Hence any monad q is such that whether x, or an obverse of x (namely x itself), be adjoined to q, always $O(xq)$. Hence, by principle II, $O(q)$. Therefore if a *single* element x exists such that $O(x)$, *every* element of Σ, as for instance q, is such that $O(q)$. Hence, by principle I, whatever collection of elements be adjoined to any element q of Σ, the resulting collection is an O-collection. Hence (by 11), since all possible collections are thus O-collections, *all* the elements of Σ are mutually equivalent. But this contradicts principle IV. Hence it is impossible that any monad x should exist such that $O(x)$. Every monad, therefore, is an E-collection.

36. It will be convenient, at this point, to restate the theorems of 21–24 in the notation for obverse elements which has now been adopted:

(1) If $O(\alpha x)$ and $O(\alpha\bar{x})$, then $O(\alpha)$ (see 21).

(2) If $O(\pi)$, then $O(\bar{\pi})$ (see 22).

(3) If $O(\beta x)$ and $O(\vartheta x)$, then $O(\beta\bar{\vartheta})$ (see 23).

(4) If $O(\delta\epsilon)$, while λ exists such that $O(\lambda e_n)$ for every element e_n of ϵ, then $O(\bar{\delta}\lambda)$ (see 24).

(5) If $O(d\epsilon)$, while λ exists such that $O(\lambda e_n)$ for every element e_n of ϵ, then $O(\bar{d}\lambda)$ (see 24).

37. Any repetitions of an element which occur in an O-collection may be stricken out, so that the element in question occurs but a single time; and the resulting collection will still be an O-collection. For suppose α to be a collection consisting wholly of the element a, repeated any multitude of times. And suppose β to be such that $O(\alpha\beta)$. From $O(\alpha\beta)$ follows by 36 (2), $O(\bar{\alpha}\bar{\beta})$. If any element, either of $\bar{\alpha}$, or of $\bar{\beta}$, be adjoined to the collection (a, β) (which consists of a, occurring only once, with β adjoined), it is plain that the resulting collection will be enlarged so as to constitute an O-collection. For $O(a\bar{a}\beta)$; and, if b_n be any element of β, and $\bar{b}_n$ the obverse of this element $O(a\beta\bar{b}_n)$. The collection $(\bar{\alpha}\bar{\beta})$, consisting entirely of complements of (a, β), is thus an O-collection. Hence $O(a\beta)$, by principle II.

38. It further follows that, if β itself is also a collection consisting solely of a repeated, and if $O(\alpha\beta)$ then the collection $O(\alpha\beta)$ which now consists solely of a repeated, can be reduced to $O(aa)$, and hence to $O(a)$. But $O(a)$ is impossible by 35. Hence no collection consisting solely of repetitions of a single element can be an O-collection. Every such collection must be an E-collection.

It still further follows that, *no element is equivalent to its own obverse.* For since $O(x\bar{x})$, it follows, by the definition of equivalence, that, if $x = \bar{x}$, $O(xx)$, Hence by 37, $O(x)$. But this, by 35, is impossible.

39. If a and b are such that the mutually obverse pairs $(\bar{a}, b)$ and $(a, \bar{b})$ are such that $O(\bar{a}bx)$ and $O(a\bar{b}x)$, then by 21 and 22, $O(\bar{a}\bar{b})$. Hence, by 34, $a = b$.

40. If any collection δ is such that $O(\delta)$ is false, so that $E(\delta)$ is true, it follows, from principle I, that if ϵ is any collection such that ϵ consists wholly of elements which belong to the collection δ, while ϵ does not include all of these elements, then $O(\epsilon)$ is false. For if $O(\epsilon)$, then, since δ may be formed from ϵ by adjoining to ϵ certain elements, principle I would require that $O(\delta)$ should be true. Whatever elements, therefore, we omit from an E-collection, δ, the remaining elements form an E-collection. Or, in other words, if a collection is an E-collection, all possible partial collections that can be formed by selecting some of its elements, and omitting others, are also E-collections, so that if ϵ is any such partial collection, $O(\epsilon)$ is false.

41. If $E(a\bar{b})$, then $a \neq b$. For if $a = b$, then, by 34, $O(a\bar{b})$, which contradicts the hypothesis $E(a\bar{b})$. And, on the other hand, if a and b are such that $a \neq b$, then $E(a\bar{b})$, for if not, then $O(a\bar{b})$, and, therefore, by 34, $a = b$, which contradicts the hypothesis $a \neq b$. Thus then, if two elements, x and y, form an E pair, the obverse of either of these elements is not equivalent to the other element; i. e., $x \neq \bar{y}$; and $\bar{x} \neq y$. Plainly, furthermore, if $E(x\bar{y})$, then $E(\bar{x}y)$, and conversely, again, if x exists such that either $E(a\bar{b}x)$ or $E(\bar{a}bx)$ is true, then, by 40, either $E(ab)$, or else $E(\bar{a}\bar{b})$ is true. But as we have just seen each of the assertions: $E(a\bar{b})$ and $E(\bar{a}\bar{b})$, implies the other, and also implies $a \neq b$. In the same way, if $E(xyz)$, then $E(xy)$, $E(yz)$ and $E(xz)$. Hence $x \neq \bar{y}$, $y \neq \bar{z}$, etc. In general, if $E(\alpha)$, and if x and y are any two of the elements of α, then, by 40, $E(xy)$, and hence $x \neq \bar{y}$. This is also immediately evident from principle I, and from the definition of obverses. For if $x = \bar{y}$, then $O(xy)$, and then any collection into which both x and y enter is, by principle I, an O-collection.

42. If $E(\alpha)$, and if x be any element whatever, then either $E(\alpha x)$ or $E(\alpha\bar{x})$ must be true. For if neither of these assertions is true, then $O(\alpha x)$ and $O(\alpha\bar{x})$; and then, by principle II, $O(\alpha)$.

43. By 35 every monad is an E-collection. Let x and y be any two elements of the system Σ. Each of these elements possesses an obverse. Since $E(x)$, by 42 either $E(xy)$ or $E(x\bar{y})$ is true; and since $E(y)$, either $E(xy)$ or $E(\bar{x}y)$ is true.

Chapter III. The F-Collections.

44. If two collections, β and η, are such that $O(\beta\bar{\eta})$, then the collections β and η stand to each other in a relation which we shall also, at pleasure, express

by the symbol $F(\beta|\eta)$, wherein the symbols β and η appear with a short vertical line between them. This symbol then, in the first place, expresses precisely the *same* facts that are expressed by the symbol $O(\beta\bar{\eta})$. That is the symbol $F(\beta|\eta)$ may at pleasure be read as the assertion: "The collection consisting of β taken together with the collection η (the collection which is the obverse of η), constitutes in its totality, an O-collection." Since, by 22 and 33, $O(\beta\bar{\eta})$ implies $O(\bar{\beta}\eta)$, and conversely, the symbol $F(\beta|\eta)$ could equally well be read: — "The collection $(\bar{\beta}, \eta)$ is an O-collection." But the symbol $F(\beta|\eta)$ is especially intended to emphasize the fact that, when $O(\bar{\beta}\eta)$, and consequently when $O(\beta\bar{\eta})$, the collections β and η stand to each other in a relation which is *mediated by the existence of their respective obverse collections.* β and η are then collections such that each, if adjoined to the obverse of the other collection, unites with that obverse, to constitute a total collection that is an O-collection. Expressing this fact with a primary reference to β and η, instead of to β and $\bar{\eta}$, or to $\bar{\beta}$ and η, the symbol $F(\beta|\eta)$ may now be read as the assertion: "The collection β forms, *with* the collection η, a *determinate F-collection.*" The special significance of the adjective *determinate* will appear below. The vertical line is intended as a sort of punctuation mark, to indicate the distinction between the two collections in question.

45. If $O(\gamma)$, and if hereupon γ be in *any* way exhaustively divided into two "partitions," that is, into two mutually exclusive collections of elements δ and ϵ, such that $O(\delta, \epsilon)$ is the same collection as γ, it is plain from the foregoing that $F(\delta|\bar{\epsilon})$. So the same O-collection makes possible various different assertions in terms of determinate F-collections. If γ is a collection of unrestricted multitude, the multitude of the possible assertions in terms of F-collections becomes also unrestricted.

46. The rule for transforming our assertions so that what are explicitly defined as O-collections shall appear in the form of explicitly designated and determinate F-collections, is consequently as follows: If the assertion $O(\gamma)$ is given, and if we are to express this as an assertion regarding some determinate F-collection, then we choose at random any partial collection δ of the elements of γ. Let $\bar{\epsilon}$ be the collection which is the obverse of the collection ϵ, where ϵ consists of all the remaining members of γ, not included in δ. Write $F(\delta|\bar{\epsilon})$, or, at pleasure, $F(\bar{\epsilon}|\delta)$. That is, put δ on one side and $\bar{\epsilon}$ on the other side of the vertical. The way in which the partial collections included in the parenthesis are placed, in so far as these two collections are merely considered with respect to their succeeding or preceding the vertical, is then capable of transposition at pleasure. The resulting expression is to be read, as above defined, and as an assertion in terms of a determinate F-collection. Instead of $F(\bar{\epsilon}|\delta)$, we can equally well write $F(\epsilon|\bar{\delta})$, or $F(\bar{\delta}|\epsilon)$.

47. The rule for the inverse operations transforming an assertion regard-

ing a determinate F-collection into an assertion regarding an O-collection is now obvious. If the assertion $F(\alpha|\beta)$ is given, then we first take the collection $\bar{\alpha}$, or, at pleasure $\bar{\beta}$; that is, we take the obverse collection corresponding that collection which stands on one chosen side of the vertical; and then we combine this obverse collection with that collection which stands on the other side of the vertical. Hereupon we write $O(\bar{\alpha}\beta)$ or $O(\beta\bar{\alpha})$, or $O(\bar{\beta}\alpha)$ or $O(\alpha\bar{\beta})$, at pleasure.

48. In case expressions such as $F(\alpha\beta x|\gamma\delta\bar{y}z)$, or similarly complex symbols appear, we read in this way: "The collections α and β, together with the element x, constitute a collection which, taken as a whole, forms an F-collection with a collection consisting of the partial collections γ and δ, and of the elements $\bar{y}$ and z." The expression just set down asserts the same as is asserted by the symbol $O(\alpha\beta.x\bar{\gamma}\delta y\bar{z})$, or as is asserted by the symbol $O(\bar{\alpha}\bar{\beta}\bar{x}\gamma\delta\bar{y}z)$.

49. If a collection, λ, consists of some finite number, n, of elements, it is one of a set or group of 2^n collections which can be formed from the given collection by first leaving that collection unchanged, and by then transforming, in every possible way, one, two, three, $\cdots$ and finally all of the n elements of λ into their respective obverses. If one of these 2^n collections is an O-collection, e. g., if the original collection λ is an O-collection; then the collection $\bar{\lambda}$, which is one of the set of 2^n collections, is also an O-collection. All of the other collections of the set are hereby required to be determinate F-collections. But in symbolizing these determinate F-collections, the arrangement of the elements is no longer wholly indifferent. One must, in each case, set upon one side of the vertical all of those elements which, in any one of the transformed collections, are obverses of elements of λ; upon the other side one must place all those elements which are identical with elements of λ. One is then to set the rearranged collection within a parenthesis, and is to write F before this parenthesis. In each case one thus asserts the same fact as is asserted by $O(\lambda)$; but does so, each time, with a different stress upon the partial collections whose relations to one another are thus pointed out. By means of the determinate F-collections, one thus analyzes, in a particular way, the various aspects of the meaning of the assertion $O(\lambda)$. Yet each one of the determinate F-collections points back, infallibly, to the same pair of O-collections; and also predetermines the constitution of all the other determinate F-collections of the same set; so that, in thus emphasizing various aspects of the meaning of the assertion $O(\lambda)$, one still never loses the power to return from one aspect, thus emphasized, to any or to all the other aspects of the same assertion. The determinate F-collections thus defined may be grouped in $(2^n-2)/2$ pairs; since, in general, if $F(\alpha|\beta)$, $F(\bar{\alpha}|\bar{\beta})$ is true. When we enumerate the set of determinate F-collections, it is sufficient to name one of each pair.

50. *Indeterminate F-collections.* It is, however, occasionally convenient to

express simply the assertion that there exists *some* O-collection $O(\bar{\kappa}\lambda)$, such that the collection (κ, λ) is precisely the same collection as a given collection η, that is, such that (κ, λ) stands for a partition of η; while we nevertheless leave it entirely undetermined *what one* of the possible partitions of the collection η it is with regard to which this assertion holds true. In this case we may write simply $F(\eta)$, a symbol which we read as the assertion: "The collection η is an (indeterminate) F-collection." An indeterminate F-collection may prove, when its determination is specified, to be any one of the determinate F-collections which correspond to the possible partitions of η. Thus, if η is the same collection as (κ, λ) and if $F(\kappa|\lambda)$ is true, $F(\eta)$ is true; while if we merely know that $F(\eta)$ is true, we know that some one of the assertions $F(\kappa|\lambda)$, $F(\kappa'|\lambda')$, etc., is true — where $(\kappa, \lambda)(\kappa', \lambda')$, etc., are various possible partitions of the single collection η. In the same way, if $F(\eta)$, some one of the assertions $O(\bar{\kappa}\lambda)$, $O(\bar{\kappa}'\lambda')$, etc., is true.

If, setting out from the assertion $O(\gamma)$, we consider some possible partition of the collection γ, say (δ, ϵ), and then, instead of writing, as above (45), $F(\delta|\bar{\epsilon})$ we write simply $F(\delta\bar{\epsilon})$, we surrender some of the information conveyed in the original assertion $O(\gamma)$, as well as in the assertion $F(\delta|\bar{\epsilon})$. For it now no longer appears what determinate F-collection corresponds to the indeterminate F-collection $F(\delta\bar{\epsilon})$; and the latter assertion tells us only that some one of the possible partitions of the collection $(\delta, \bar{\epsilon})$ is such that, if it is made (e. g., the partition $(\delta_1, \delta_2, \bar{\epsilon}_1, \bar{\epsilon}_2)$, wherein (δ_1, δ_2) is the same collection as δ, and $(\bar{\epsilon}_1, \bar{\epsilon}_2)$ the same collection as $\bar{\epsilon}$) — then $F(\delta_1\epsilon_1|\delta_2\bar{\epsilon}_2)$, so that $O(\delta_1\bar{\epsilon}_1\bar{\delta}_2\epsilon_2)$. It is plain that the indeterminate F-collections occur in pairs. If $F(\alpha)$, then $F(\bar{\alpha})$.

51. An example will serve to distinguish more clearly the kinds of information conveyed by the assertion that a collection is a determinate and by the assertion that this collection is an indeterminate F-collection. Let $F(ab|cd)$ be true. This is equivalent to asserting $O(ab\bar{c}\bar{d})$. Any one of the possible collections which can be formed by transforming (a, b, c, d) through the substitution of the obverse of one or of more of its elements, is then also a determinate F-collection. Thus the assertion $F(ab|cd)$ requires:

For the collection $(a, b, c, \bar{d})$ the assertion $F(ab\bar{d}|c)$;
For the collection $(a, b, \bar{c}, d)$ the assertion $F(ab\bar{c}|d)$;
For the collection $(a, \bar{b}, c, \bar{d})$ the assertion $F(a\bar{d}|c\bar{b})$;
For the collection $(a, \bar{b}, \bar{c}, \bar{d})$ the assertion $F(a\bar{c}\bar{d}|\bar{b})$;

and so on; while each of these assertions implies, and is implied by $O(ab\bar{c}\bar{d})$.

The assertion $F(abcd)$ does not necessarily imply any one of the foregoing determinate assertions. It indicates that, of the $2^n - 1$ or 15 possible collections other than (a, b, c, d) in the set of collections producible from (a, b, c, d) through the substitution of obverses, some one, and consequently (since $O(\alpha)$ implies $O(\bar{\alpha})$), some *pair* of collections, must be O-collections. It would

depend upon this pair to determine how the vertical lines which define the determinate F-collections ought to be distributed in each of the various cases which hereupon arise. Some one of these possible distributions express the truth in each case; but the assertion $F(abcd)$ does not tell us which one this is. Thus, if $F(abcd)$, and if we then consider the collection $(a, \bar{b}, c, \bar{d})$, it is possible that:

$F(a|\bar{b}c\bar{d})$, i. e., $O(ab\bar{c}d)$, and $O(\bar{a}\bar{b}cd)$; or $F(a\bar{b}c|\bar{d})$, i. e., $O(\bar{a}bcd)$, and $O(\bar{a}b\bar{c}\bar{d})$; or $F(ad|\bar{b}c)$, i. e., $O(\bar{a}\bar{b}cd)$ and $O(ab\bar{c}\bar{d})$, and so on for all the other cases.

The assertion $F(abcd)$ requires some pair of these alternative O-assertions or some corresponding pair of the F-assertions, to be true, but does not specify what pair in any of the cases in question.

The indeterminate F-collections, like the O-collections, are perfectly symmetrical. In case of a pair (a, b), the alternative pairs of assertions:—

$$\left\{ \begin{array}{c|c} (1) & \begin{array}{c} O(ab) \\ O(\bar{a}\bar{b}) \end{array} \end{array} \right| \left. \begin{array}{c|c} (2) & \begin{array}{c} O(a\bar{b}) \\ O(\bar{a}b) \end{array} \end{array} \right\}$$

are such that if the pair (1) of assertions are both true, the assertions (2) are both of them false. Hence $F(ab)$ can mean only that $O(\bar{a}b)$ and $O(a\bar{b})$ are both of them true; while $F(a\bar{b})$ means that both $O(ab)$ and $O(\bar{a}\bar{b})$ are true. Hence, in case of pairs of elements, the distinction between determinate and indeterminate F-collections vanishes: and the assertion $F(ab)$ is perfectly determinate.

Elementary properties of the F-collections: operations and transformations.

52. A number of elementary properties of F-collections, and a survey of certain ways in which they may be transformed, may now be readily obtained from the already established properties of the O-collections.

(1) If $F(xy)$ then, since $O(x\bar{y})$, $x = y$. See 34.

(2) If $F(ax|b)$ and $F(bx|a)$, then $a = b$. For $O(a\bar{b}x)$ and $O(\bar{a}bx)$. See 39.

(3) If $F(\eta)$, then $F(\eta\gamma)$ where γ is any collection whatever (Principle I).

(4) If $F(\eta)$, then $F(\bar{\eta})$, as was already observed in connection with the definition of the F-collections. If $F(\beta|\eta)$, then $F(\bar{\beta}|\bar{\eta})$.

(5) If $F(\eta|x)$ and $F(\eta|\bar{x})$, then $O(\eta)$, since $O(\eta\bar{x})$ and $O(\eta x)$. The question may then arise whether $F(\eta)$ is in a given case, also true. It is here first obvious that $F(\eta)$ does not follow from $F(\eta|x)$ and $F(\eta|\bar{x})$. For instance, if $F(ab|x)$ and $F(ab|\bar{x})$, then $O(ab)$. But, if $O(ab)$, $a = \bar{b}$; while if $F(ab)$, $a = b$ (by (1) of the present paragraph). By 38, however, if $O(ab)$, it is impossible that $a = b$. Hence, if $F(ab|x)$ and $F(ab|\bar{x})$,

$F(ab)$ is false. On the other hand, if $O(\alpha)$, then by adjunction, $O(\alpha x\bar{y})$ and $O(\alpha xy)$. Hence, in this case, $F(\alpha x|y)$ and $F(\alpha x|\bar{y})$. But in this case also since $O(\alpha)$, $O(\alpha\bar{x})$. And so $F(\alpha|x)$, i. e. $F(\alpha x)$, is likewise true. If $F(\eta|x)$ and $F(\eta|\bar{x})$, $F(\eta)$ is accordingly possible; but does not follow herefrom.

53. Any determinate F-collection remains such when transformed according to the following rule: *Substitute* for any element, or for any collection, which stands upon one side of the vertical, *the obverse* of that element or collection, *transfer* the obverse in question to the other side of the vertical, being careful to retain, as the result of the transfer, at least one element on each side of the vertical. Thus, if $F(x\beta|y\delta)$, then $F(\beta|\bar{x}y\delta)$, $F(\bar{y}\beta x|\delta)$, $F(x|\bar{\beta}y\delta)$, $F(\beta\bar{y}\bar{\delta}|\bar{x})$, etc., are all of them true. This is obvious, because all these expressions mean the same as $O(\bar{x}\bar{\beta}y\delta)$ or as $O(x\beta\bar{y}\bar{\delta})$.

This is called transformation by transfer. If the elements are transferred by this rule except that *all* the elements are permitted at the end to stand upon *one* side of the vertical, the vertical can then be omitted; but the resulting collection must be regarded as thus transformed into an O-collection.

54. If $F(x\beta|\bar{x}\gamma)$ then $F(\beta|\bar{x}\gamma)$ and $F(x\beta|\gamma)$. If $F(x\beta|\bar{\beta}\gamma)$, then $F(x\beta|\gamma)$ and $F(x|\bar{\beta}\gamma)$. That is, if the obverse of an element or of a collection, which stands on *one* side of the vertical, itself stands on the *other* side of the vertical, then either of the two mutually obverse collections or elements may be stricken out (by 37). For if $F(x\beta|\bar{x}\gamma)$, then $O(x\beta x\bar{\gamma})$, and so $O(\beta x\bar{\gamma})$. Hence $F(\beta x|\gamma)$ and $F(\beta|\bar{x}\gamma)$. This is called a transformation by means of the omission of superfluous obverses; and the procedure obviously applies to collections as well as to elements.

55. If $F(\alpha|\beta)$, and if all the elements of α are mutually equivalent, and all the elements of β are mutually equivalent, then all the elements of the collection are mutually equivalent.

For $O(\alpha\bar{\beta})$. Let a be one of the elements of α and b of β. Then, by 36, $O(\alpha\bar{\beta})$ reduces to $O(a\bar{b})$, whence follows, by 34, $a = b$.

It is now obvious that any repetitions of an element which occur upon one side of the vertical in a determinate F-collection may be stricken out; and also that, by virtue of principle I, any element may be added to that collection which stands upon either side of the vertical, so that, if $F(\alpha|\beta)$, $F(\alpha y|\beta)$ and $F(\alpha|\beta y)$, where y is any element.

Elimination-theorems for F-collections.

56. If two collections β and δ are such that there exists a collection π such that $F(\delta|\pi)$, while, for every member y_r of the collection π, $F(y_r|\beta)$, then $F(\beta|\delta)$.

For, if $F(\delta|\pi)$, then $O(\delta\bar{\pi})$. And if $F(y_r|\beta)$, then, for every member

$\overline{y}_r$ of the collection $\overline{\pi}$ it is true that $O(\overline{y}_r\beta)$. Substituting for the symbol $\overline{\pi}$ the symbol ϵ, and for $\overline{y}_r$ (the representative symbol for any member of $\overline{\pi}$), the symbol x_r (as the representative symbol for any member of the collection now called ϵ), we have δ and β such that there exists a collection ϵ such that $O(\delta\epsilon)$, while, for every member x_r of ϵ, $O(x_r\beta)$. Hence, by 36, and by virtue of the properties of O-collections pointed out in 24, $O(\delta\overline{\beta})$. Hence $F(\delta|\beta)$ or $F(\beta|\delta)$.

57. In case the collection δ reduces to the single element d, the theorem assumes the following form:

If any collection β forms a determinate F-collection with every member of π, separately considered, while the collection π taken as an entirety, is such as to form a determinate F-collection with an element d, then β forms a determinate F-collection with d.

This theorem permits the elimination of π, in case $F(d|\pi)$, and in case the set of determinate F-collections $F(y_r|\beta)$ is given, where y_r is a variable for which every element of π may separately be substituted.

58. If π reduces to a single member y, we have the result of 56 reduced to the form:

If two collections β and δ are such that there exists an element y such that $F(y|\beta)$ and $F(y|\delta)$, then $F(\beta|\delta)$.

This last result furnishes a means for the direct elimination of an element y common to two determinate F-collections, in case y stands alone, on one side of the vertical, in each collection. Here too we deal with a type of *transitivity* whose consequences are of great importance.

59. If $F(x\beta|y\gamma)$, and $F(y\beta|z\gamma)$, then $F(x\beta|z\gamma)$. For, by transfer (53), from $F(x\beta|y\gamma)$ follows $F(x\beta\overline{\gamma}|y)$. And from $F(y\beta|z\gamma)$ follows $F(\overline{\beta}z\gamma|y)$. Hence, by 58, we can eliminate y, and thus we obtain $F(x\beta\overline{\gamma}|\overline{\beta}z\gamma)$. By 54, we may hereupon transform this F-collection by striking out $\overline{\beta}$ from the right side of the vertical (since β occurs on the left side), and $\overline{\gamma}$ from the left side (since γ occurs on the right side). We thus obtain $F(x\beta|z\gamma)$, which was to be proved.

The transformations and the type of elimination here used are typical of the methods which are to be employed in considering and in combining F-collections. These methods correspond to the adjunctions and eliminations already used in case of O-collections.

60. The result of 59 is the principal theorem relating to the transitivity of the relations involved in F-collections. Its importance justifies a proof directly in terms of the properties of O-collections.

If, namely, $F(x\beta|y\gamma)$, then $O(x\beta\overline{y}\overline{\gamma})$. From this follows, by adjunction, $O(x\beta\overline{\gamma}\overline{y}\overline{z})$. If $F(y\beta|z\gamma)$, then $O(y\beta\overline{z}\overline{\gamma})$. From this follows, by adjunction, $O(x\beta\overline{\gamma}y\overline{z})$. From the two O-pentads, thus formed through adjunction, follows, by principle II, $O(x\beta\overline{\gamma}\overline{z})$. Hence $F(x\beta|\gamma z)$.

61. The hypothesis $F(x\beta|y\gamma)$ sets x in a determinate dyadic relation R to y; and this relation is unsymmetrical, since y does not stand in this relation to x. The hypothesis $F(y\beta|z\gamma)$ sets y in this same relation R to z. The conclusion sets x in the relation R to z. The relation in question is therefore transitive. But the unsymmetrical transitive dyadic relation in question has been entirely derived from the wholly symmetrical relations defined by the O-collections.

62. If β, γ, and δ are given collections, and if y and e are given elements, such that $F(y\beta|\delta)$ and $F(e\delta|\beta)$, while $F(e|y\gamma)$, then $F(\gamma\delta|\beta)$.

For by transfer (53), from $F(y\beta|\delta)$ and $F(e|y\gamma)$ follow the two assertions $F(y|\delta\bar{\beta})$ and $F(y|e\bar{\gamma})$. By 58, it follows that $F(\delta\bar{\beta}|e\bar{\gamma})$; whence follows again by transfer $F(\bar{\beta}\gamma\delta|e)$. From $F(e\delta|\beta)$ follows, by transfer, $F(e|\beta\bar{\delta})$. From this and $F(\delta\bar{\beta}\gamma|e)$ follows, by 58, $F(\beta\bar{\delta}|\gamma\delta\bar{\beta})$. By omission of superfluous obverses (54), we obtain $F(\beta\bar{\delta}|\gamma)$ and $F(\beta|\gamma\delta)$.

63. If $F(x|\beta)$ and $F(y|\beta)$, while $F(xy|\delta)$ then $F(\delta|\beta)$.

This follows directly from 56, in case the collection π of that theorem reduces to the pair (x, y).

64. If α, δ, ϵ are collections, and b and c are elements, such that (1) $F(b\alpha|\delta)$ and (2) $F(c\alpha|\delta)$, while (3) $F(bc|\epsilon)$, then $F(\delta|\alpha\epsilon)$.

For by (1) $F(b\alpha|\delta)$ while by (2), $F(\alpha|\delta\bar{c})$. Let $b = y$, and $c = \bar{e}$. Then $F(y\alpha|\delta)$, while $F(\alpha|\delta e)$; and, by (3), $F(y|\epsilon e)$, i. e., $F(e|y\bar{\epsilon})$.

Hence, by 62, $F(\bar{\epsilon}\delta|\alpha)$. Whence follows $F(\alpha\epsilon|\delta)$. That is, if two collections α and δ are such that if either b or c be separately adjoined to α, the resulting collection forms an F-collection with δ, and if the pair (b, c) forms an F-collection with ϵ, then if ϵ itself be adjoined to α, the resulting collection forms an F-collection with δ.

Chapter IV. F-triads, mediators and antecedents.

65. What elements exist in the system Σ we have as yet but very imperfectly investigated. Yet before we proceed to this investigation, it will prove convenient to outline the general character of the order which is possible in the system Σ, so far as we have yet developed this order. The fact that given elements do or do not belong to a certain O-collection, or do or do not constitute an O-collection, is one which appears directly to establish no sort of order amongst the elements of Σ. The relation in which various elements stand to one another when they belong to the same O-collection, is so far absolutely *symmetrical*, and nothing can be said of one member of such a collection which is not asserted of everyone of the others, so far as this collection is concerned.

But the fact that every element of Σ possesses an obverse, enables one to establish relations between certain elements, or sets of elements, relations which are due to the further fact that given elements may enter into O-collections with the

obverses of certain other elements. The consequence of this is that F-collections are definable. Let us provisionally assume that a large variety of F-collections exist in Σ.

66. F-collections, if indeterminate, are, like O-collections, of a wholly symmetrical structure, and their members are so far undistinguished from one another. But the members of a determinate F-collection are no longer so symmetrically disposed. It is indeed true that the order in which the elements in the set of an F-collection on each side of the vertical are considered, is indifferent. Nor does it make any difference which set is written before or after the vertical. But if $F(\beta|\gamma)$, then, in general, any element x of β is related to any element y of γ in a way which is *not* reciprocated. For x is related to y as that element which, in combination with some collection β' of companion elements, forms an F-collection with y, when y is combined with a collection γ' of elements (β' being the collection of the other elements of β besides x, γ' being the collection of the other elements of γ besides y). This relation, if read in the other direction, changes, in general, its character, and so is an unsymmetrical relation.

But, as 59 has shown us, this unsymmetrical dyadic relationship is *transitive*. In terms of this relation certain sets of the elements may be *ordered* and so arranged in series like points on a line.

67. The fundamental form of such series becomes manifest if we pay attention to those cases of the much more general theorems regarding F-collections, which appear as special results if we consider only *triads* of elements.

In this case if, for example, $F(ac|b)$ or $F(ca|b)$, that is, if $O(ac\bar{b})$, we shall call b "the *mediator* between a and c," or, where that is more convenient, the "mediator of the pair (a, c)." The "mediator" of a pair is accordingly an element whose obverse forms an O-triad when adjoined to that pair. If $O(pqr)$, then $F(pq|\bar{r})$, so that the obverse of any member of an O-triad is the mediator of the pair formed by the other elements of that triad.

The relation of the mediator to the elements which it mediates may be treated at pleasure either as a triadic or as a dyadic relation. In order to treat it as a dyadic relation we may take account of the fact that, if $F(ac|b)$, b is in a certain relation to a *with respect to* c, and is in a certain relation to c *with respect to* a. This aspect of the matter may become especially important in case we deal with a number of triads in which the mediators are any elements whatever, while all the pairs mediated have a common element y. Such pairs appear if $F(ym|q)$, $F(yn|r)$, $F(yo|s)$, etc. In all such cases q has the same relation to m that r has to n, and that s has to o, etc., since q is in a given relation to m *with respect to* y; and the same holds true in the other instances in question.

68. If $F(yq|p)$, we may, whenever that is convenient, first select one element of the pair (q, y), say the element y, and thereupon say that p is in a relation

to q which we shall call the relation: "*antecedent of q with respect to y*." The element y we may hereupon call the "*origin*" from which the relation is defined or reckoned. Equally, if we choose the element q as our origin, we can say that p is in the relation to y of being "an antecedent of y with respect to q". We shall symbolize the assertion "p is antecedent of q with respect to y," by the expression $p \prec_y q$ or $q \succ_y p$; and we can also at pleasure read either of these expressions thus: "q is a consequent of p with respect to y." This expression means precisely the same as the expression $F(yq|p)$, or as the expression $O(yq\bar{p})$. Expressing the facts in the new way, as involving the relation of an antecedent to a consequent, has merely the advantage of bringing out certain aspects of the situation which will be conveniently expressible in terms of a dyadic relation — a relation which, as already pointed out, will prove to be unsymmetrical and transitive and, therefore, useful for the definition of serial order amongst certain specially selected elements of Σ. The equivalence of meaning of the three expressions: $O(qy\bar{p})$, $F(qy|p)$, and $p \prec_y q$, enables us at once to see how superficial is the difference between symmetrical and unsymmetrical relations. All that any one of these three expressions asserts is that the two perfectly symmetrical O-collections $O(qyx)$ and $O(xp)$ (where $x = \bar{p}$), both exist as collections of the elements of Σ.

69. It is of course expressly true that the relation of antecedent to consequent, or of consequent to antecedent, has meaning *only* with reference to a given origin. It is this origin which gives "*sense*" to the pair (pq) in the expression $p \prec_y q$. On the other hand, if a question arises as to whether the "sense," or asymmetry, of the dyadic relation here in question, is a fundamental fact, or is unanalyzable — this question is answered in advance by our derivation of the whole asymmetry from the perfectly symmetrical properties which characterize the various members of any O-collection.

The system Σ, as we shall hereafter see, includes elements whose relations are precisely the ones which are of the most fundamental importance in all the exact sciences. The customary procedure of these sciences may be said, in the main, to involve the definition of these relations in terms of the relation of antecedent and consequent. Wherever a linear series is in question, wherever an origin of coördinates is employed, wherever "cause and effect," "ground and consequence," orientation in space or direction of tendency in time are in question, the dyadic asymmetrical relations involved are essentially the same as the relation here symbolized by $p \prec_y q$.

This expression, then, is due to certain of our best established practical instincts and to some of our best fixed intellectual habits. Yet it is not the only expression for the relations involved. It is in several respects inferior to the more direct expression in terms of O-relations. The range of its efficacy as an expression will become clearer hereafter. When, in fact, we attempt to

describe the relations of the system Σ merely in terms of the antecedent-consequent relation, we not only limit ourselves to an arbitrary choice of origin, but miss the power to survey at a glance relations of more than a dyadic, or triadic character.

Properties of the relation of mediator, antecedent, and consequent.

70. If $F(ab|c)$, and $a = b$, then $c = a = b$. For this is a special case of the principle proved in 55.* Furthermore, if $p \prec_y q$ and $q \prec_y p$, then $p = q$, since this is but another expression of 52 (2). Consequently, if $p \neq q$, and $p \prec_y q$, then $q \prec_y p$ is false. The relation of antecedent to consequent, if it obtains at all between a pair of non-equivalent elements, is therefore in that case inconvertible; hence the relation becomes totally asymmetrical so soon as it is confined to pairs of non-equivalent elements.

71. If $F(a|by)$ and $F(b|cy)$, then $F(a|cy)$ and $F(b|ac)$.

If $a \prec_y b$ and $b \prec_y c$, then $a \prec_y c$; and $b \prec_a c$.

The proof is as in 59. Thus from $F(b|a\bar{y})$, and $F(b|cy)$ follows $F(a\bar{y}|cy)$ and $F(a|cy)$. From $F(y|a\bar{b})$ and $F(y|b\bar{c})$ follows $F(\bar{ab}|b\bar{c})$; and consequently $F(b|ac)$. The form of the theorem stated on the right is simply a direct translation into the symbolism of the relation of antecedent and consequent. The latter relation is thus shown to be transitive.

72. If $F(x|ab)$ and $F(x|bc)$, while $F(b|ac)$, then $x = b$. For since $F(x|bc)$ while $F(b|ca)$, we have $F(c|x\bar{b})$ and $F(c|b\bar{a})$. Whence follows $F(x|b\bar{a})$; that is $F(b|ax)$. From $F(x|ab)$ and $F(b|ax)$ follows, by 52 (2) $x = b$.

In other words, if $x \neq b$, and b is a mediator of the pair (a, c), it is impossible that x should be at once a mediator of the two pairs (a, b) and (b, c).

73. If $F(x|ab)$ and $F(y|ab)$ and $F(d|xy)$, then $F(d|ab)$.

If $x \prec_a b$ and $y \prec_a b$ and also $d \prec_x y$, then $d \prec_a b$.

This follows directly from 63 and expresses the fact that a mediator of two elements which, with respect to a given origin, are antecedents of the same element, is itself an antecedent of that element with respect to the same origin.

74. If $F(yb|d)$, $F(yc|e)$ and $F(ed|b)$, then $F(cd|b)$.

If $d \prec_y b$ and $e \prec_y c$, while $b \prec_d e$, then $b \prec_d c$.

This is a special case of the proposition proved in 62. If two pairs, (b, y) and (c, y), have a common element y, and if each pair forms a determinate F-triad when a term d (in one case), or e (in the other case) is set on the opposite side of the vertical, and if the member b of the one pair is a mediator of e and d, while d is the third member of the F-triad in which b occurs, then b is a mediator of d and of c, where c is the remaining member of the other pair (c, y).

* This theorem is used by KEMPE as a fundamental principle in defining the relation here called that of Mediator.

Otherwise: If, with respect to a common origin y, the elements d and e are, respectively, antecedents of the elements b and c, so that d is antecedent of b and e of c, then, if b is a mediator of d and e, b is also a mediator of c and d.

75. If $F(ab|d)$, $F(ac|d)$ and $F(bc|e)$, then $F(ae|d)$.

If $d \prec_a b$ and $d \prec_a c$ and $e \prec_b c$, then $d \prec_a e$.

This follows from 64.

That is, whatever element is mediator of two pairs which have an element in common, is mediator of the pair composed of the common element and of any mediator of the pair, formed by the elements which, belonging to the original pairs are not common to them. Otherwise, whatever element is, with respect to a given origin, an antecedent of each of a pair of elements, is antecedent of any mediator of this pair.

76. Any element a is a mediator between any element y and the obverse of y; and any element y is a mediator between itself and any other element a.

For $O(\bar{a}y\bar{y})$ and $O(ay\bar{y})$.

77. For any origin y, it is true, that y is an antecedent of itself and of every other element, including $\bar{y}$; while any element a is an antecedent of itself and of $\bar{y}$.

For, $O(\bar{y}ya)$, $O(\bar{y}yy)$ and $O(\bar{y}y\bar{y})$, while $O(\bar{a}ay)$ and $O(\bar{a}y\bar{y})$.

78. If $a \prec_y b$, then $b \prec_y \bar{a}$.

For $O(\bar{a}by)$, and hence $F(\bar{b}|\bar{a}y)$.

79. Whatever element is, with respect to a given origin y, an antecedent of every member of an O-collection except one, is also an antecedent of the obverse of this excepted element.

Let π be any collection, and e such an element that $O(\pi e)$. Then, by this hypothesis $F(\pi|\bar{e})$. Let q be an element such that $F(q|x_r y)$ is true of every element x_r of π so that $q \prec_y x_r$. Then, by transfer, $F(q\bar{y}|x_r)$. Hence the collection $(q, \bar{y})$ forms a determinate F-collection, when set on one side of the vertical, with *any* element of the determinate F-collection $F(\pi|\bar{e})$ on the other side of the vertical, except the element $\bar{e}$. Hence, by 57, it is *also* true that $F(q\bar{y}|\bar{e})$. Hence $F(q|y\bar{e})$. Hence $q \prec_y \bar{e}$.

80. If, with respect to y, q is an antecedent of *every* member of a given O-collection, then q is equivalent to y. For q is in any case an antecedent, with respect to the origin y, of every member of the collection except any member e. Hence $q \prec_y \bar{e}$. But, by the present hypothesis, $q \prec_y e$ is *also* true. Hence $O(\bar{q}ye)$ and $O(\bar{q}y\bar{e})$. Hence $O(\bar{q}y)$. Hence $q = y$.

81. Whatever element is, with respect to a given origin, a consequent of every member of an O-collection except one, is also a consequent of the obverse of this excepted element.

Suppose π such that $F(\pi|\bar{e})$ as before. And suppose that, for every element x_r of π, $x_r \prec_y q$, so that $F(x_r|yq)$. By the same reasoning as that of the last theorem $F(\bar{e}|qy)$; and hence $\bar{e} \prec_y q$.

82. If q is a consequent of every member of the collection, then it is also true that $e \prec_y q$; whence $O(qye)$ and $O(qy\bar{e})$. Hence $O(yq)$, and consequently $q = \bar{y}$.

CHAPTER V. THE EXISTENCE OF ELEMENTS. CHAINS. RESULTANTS. CONJUGATE PAIRS OF RESULTANTS. THE ALGEBRA OF LOGIC.

83. In discussing, in the foregoing, the most characteristic relations of the elements of Σ, we have tacitly and provisionally assumed that elements may exist in sufficient variety to exemplify these relations. Our existential principles have so far been used mainly to establish the existence of pairs of obverse and of non-equivalent elements. We must now proceed to survey, more in detail, the actual structure of the system Σ.

84. By principles III and IV, the system Σ contains a pair of non-equivalent elements, say (xy). By 26, the pair $(\bar{x}, \bar{y})$ also exists. It is however in so far possible that $x = \bar{y}$, and hence that $y = \bar{x}$. In that case the members of the pair (x, y) become respectively equivalent to the members of the pair $(\bar{y}, \bar{x})$, or, again, to the members of the pair $(x, \bar{x})$. Were the system Σ to consist *merely* of the single pair of mutually non-equivalent and mutually obverse elements, $(x, \bar{x})$, principles I, II, III, IV, and VI would all of them be true of the system as thus restricted. For $O(x\bar{x})$; and if, to this collection, we add any collection γ, consisting either of x, or of $\bar{x}$, repeated any multitude of times, or again consisting both of x and of $\bar{x}$, in any combination, each occuring any multitude of times — in any case $O(x\bar{x}\gamma)$. Hence principle I is satisfied. On the other hand since $O(x\bar{x})$, while, by 35, 37, $E(xx)$, and $E(\bar{x}\bar{x})$ — the non-equivalence of x and $\bar{x}$ can be readily established, without taking account of any collections except those into which x and $\bar{x}$ either jointly or severally enter. In order that a collection β shall consist altogether of complements of δ, while $O(\beta)$ is true, it is necessary, in case Σ contains *only* x and $\bar{x}$, that β should include both x and $\bar{x}$; while δ (which, by hypothesis is such that $O(\delta b_n)$ for every element b_n, of β), must then *also* include both x and $\bar{x}$. Hence if $O(\beta)$, $O(\delta)$ follows. Hence principle II would hold true if the system Σ consisted only of the pair $(x, \bar{x})$. Principles III and IV would obviously hold true of the same system. And considered with reference to x, its obverse, $\bar{x}$, is an element satisfying the requirements of principle VI; while the same holds true of x when it is considered with reference to $\bar{x}$. All the principles except V would therefore be satisfied if the system Σ consisted of the single pair $(x, \bar{x})$.

Chains of elements, defined by recurrence.

85. Principle V, however, is not satisfied by the existence of a single pair of elements, such as $(x, \bar{x})$. For, if the pair $(x, \bar{x})$ exists, then (since $x \neq \bar{x}$ by

38), it follows that principle V demands the existence of r_1 such that $E(r_1 x)$ and $E(r_1 \bar{x})$. These two assertions require that $r_1 \neq \bar{x}$ and $r_1 \neq x$ (by 41). Since $r_1 \neq \bar{x}$, there exists an element which we will next symbolize by r_2, and which, by virtue of principle V, is such that $E(r_1 \bar{r}_2)$, $E(\bar{r}_2 \bar{x})$, and $O(r_1 \bar{r}_2 \bar{x})$. The obverse of $\bar{r}_2$ also exists. Since $E(r_1 \bar{r}_2)$ it follows by 41, that $r_2 \neq r_1$; and for the precisely analogous reason, $r_2 \neq \bar{x}$. At the same time $F(r_2 | r_1 \bar{x})$. Since $r_2 \neq \bar{x}$, there also exists, by principle V, an element $\bar{r}_3$, such that $E(\bar{r}_3 r_2)$ and $E(\bar{r}_3 \bar{x})$ while $O(r_2 \bar{r}_3 \bar{x})$. The obverse r_3 of $\bar{r}_3$, therefore, also exists; and is such that $r_3 \neq r_2$, $r_3 \neq \bar{x}$, and $F(r_3 | r_2 \bar{x})$. The procedure whereby the elements r_1, r_2, and r_3 have been defined is obviously a recurrent one. Repeatedly applied it defines a chain of elements: $(r_1, r_2, \cdots, r_n, \cdots)$, whereof any finite number n may be at pleasure defined in a determinate order. These elements, together with their obverses $(\bar{r}_1, \bar{r}_2, \cdots, \bar{r}_n, \cdots)$, all exist in Σ, and possess the following properties:

(1) $$x \neq r_1;\ r_1 \neq r_2;\ r_2 \neq r_3;\ \cdots;\ r_{n-1} \neq r_n;\ r_n \neq \bar{x}.$$

(2) $$F(r_1 | x\bar{x}),\ F(r_2 | r_1 \bar{x}),\ F(r_3 | r_2 \bar{x}),\ \cdots,\ F(r_n | r_{n-1} \bar{x}).$$

(3) $$F(\bar{r}_1 | x\bar{x}),\ F(\bar{r}_2 | \bar{r}_1 x),\ F(\bar{r}_3 | \bar{r}_2 x),\ \cdots,\ F(\bar{r}_n | \bar{r}_{n-1} x).$$

(4) $$\bar{x} \neq \bar{r}_1;\ \bar{r}_1 \neq \bar{r}_2;\ \bar{r}_2 \neq \bar{r}_3;\ \cdots;\ \bar{r}_{n-1} \neq \bar{r}_n;\ \bar{r}_n \neq x.$$

The expressions (3) and (4) follow directly from the truth of (1) and of (2), by the definitions of the F-collections, and of the relations of equivalent and of obverse elements, and by 30, 36(2), and 52(4).

From these conditions it further follows that, if m and n are (for the moment) viewed, not as symbols for elements of Σ, but as purely numerical marks or indices, serving to distinguish the ordinal positions of different members of the chain $(r_1, r_2, \cdots, r_m, \cdots, r_n, \cdots)$, and if m, in the series of natural numbers, precedes n, then:

(5) $$F(r_n | r_m \bar{x});\ \text{and}\ r_n \neq r_m.$$

(6) $$F(\bar{r}_n | \bar{r}_m x);\ \text{and}\ \bar{r}_n \neq \bar{r}_m.$$

For, by the laws of the construction of the chain, $F(r_{m+1} | r_m \bar{x})$, and $F(r_{m+2} | r_{m+1} \bar{x})$. Hence $F(r_{m+2} | r_m \bar{x})$, by 71. In the same way, since $F(r_{m+3} | r_{m+2} \bar{x})$, there follows $F(r_{m+3} | r_m \bar{x})$. And the same process of elimination can be repeated any number of times, so that, in fine, $F(r_n | r_m \bar{x})$. Meanwhile, by the principle proved in 72, $r_{m+2} \neq r_m$. In fact, if we suppose $r_{m+2} = r_m$, we have $F(r_{m+1} | r_m \bar{x})$ by construction, as well as $F(r_{m+2} | r_{m+1} \bar{x})$. By the substitution of equivalents we therefore obtain:

$$F(r_{m+1} | r_m \bar{x}) \qquad \text{and} \qquad F(r_m | r_{m+1} \bar{x}).$$

Whence follows, by 52 (2), $r_{m+1} = r_m$, which contradicts the conditions stated in the expressions (1). Hence $r_{m+2} \neq r_m$. In the same way we can prove that $r_{m+3} \neq r_m$; and in general, $r_n \neq r_m$, where n is any ordinal number that follows m. Hence the expressions (5) are proved true; and the expressions (6) follow therefrom by considering the obverses of the elements occurring in (5). All the members of the chain $(r_1, r_2, \cdots, r_n, \cdots)$ are such that $F(r_n | r_1 \bar{x})$.

To sum up: The requirements of principle V include the assertion that, since the pair $(x, \bar{x})$ exists there must also exist the two distinct chains of elements:

$$(r_1, r_2, \cdots, r_n, \cdots) \qquad \text{and} \qquad (\bar{r}_1, \bar{r}_2, \cdots, \bar{r}_n, \cdots).$$

No two of the elements of either one of these chains are mutually equivalent. If any element of the first chain, as r_m, were equivalent to a member of the second chain, say $\bar{r}_n$, we should have, since, by the foregoing:

$$F(r_m | r_1 \bar{x}) \qquad \text{and} \qquad F(\bar{r}_n | \bar{r}_1 x),$$

the consequence: $F(r_m | r_1 \bar{x})$ and $F(r_m | \bar{r}_1 x)$. By elimination would follow $F(r_1 \bar{x} | \bar{r}_1 x)$; whence would follow $O(r_1 x)$ and so $r_1 = x$, a result which is rendered impossible by the conditions that define r_1. Hence no element of either chain can be equivalent to any member either of the same chain, or of the other chain; and all the elements of both chains are non-equivalent both to x and to $\bar{x}$. Since, if any element r_n exists in the first chain, r_{n+1} also exists (while $\bar{r}_{n+1}$ also belongs to the other chain), the two chains contain each an infinite number of non-equivalent elements. *The system* Σ *consequently contains an infinite number of mutually non-equivalent elements.*

86. Herewith, however, the requirements of principle V are by no means exhausted. For since no two elements of either chain are mutually equivalent, any two successive elements of each chain, as, for instance, r_m and r_{m+1}, are such that, by principle V, $\bar{s}_1$ exists such that $E(r_m \bar{s}_1)$ and $E(r_{m+1} \bar{s}_1)$, while $O(r_m r_{m+1} \bar{s}_1)$. In this case $F(s_1 | r_m r_{m+1})$, while $s_1 \neq r_m$ and $s_1 \neq r_{m+1}$. A new recurrent process is thus defined, a process which can be employed to define a chain $(s_1, s_2, \cdots, s_j, \cdots)$ where the subscripts again have the significance of the ordinal numbers.

By 72, no two elements of this chain can be mutually equivalent. Every member s_j of this new chain must be non-equivalent to any member of the chain $(r_1, r_2, \cdots, r_m, \cdots)$, as well as to any member of any chain $(s'_1, s'_2 \cdots, s'_k, \cdots)$, or $(t_1, t_2, \cdots, t_h, \cdots)$, such as can be formed by taking account of those pairs: (r_1, r_2), (r_3, r_4), etc., which are different from the pair (r_{m+1}, r_m), and by applying principle V recurrently to them. In general, if $F(e | r_m r_{m+1})$, and $F(g | r_n r_{n+1})$, where m and n are different subscript numbers, while m precedes n in the ordinal series, we can express the relations already considered by taking, if we choose, r_1 as an origin, and by employing the relation of antecedent and consequent. In this case we can write, upon the basis of 85 (5),

$$r_m \prec_{r_1} r_{m+1},$$

$$F(e \mid r_m r_{m+1});\ e \neq r_m;\ e \neq r_{m-1}.$$

Whence follows $e \prec_{r_1} r_{m+1}$.

Furthermore: $r_{m+1} \prec_{r_1} r_n$ by 85 (5),

$$r_n \prec_{r_1} g,\ r_n \neq g.$$

And so, if $e = g$, we have:

$$g \prec_{r_1} r_{m+1} \prec_{r_1} r_n \prec_{r_1} g.$$

This, however, is impossible. Thus no member of any one of the new chains is equivalent to any member of any other of these new chains, or of the original chains. Since an infinite number of different pairs of non-equivalent members exist in the chain $(r_1, r_2, \cdots, r_n, \cdots)$, it follows that an infinite number of new chains can be constructed upon the basis of these various pairs. The process of forming such chains is itself recurrent.

In addition to the chains thus far defined, other chains of elements exist in Σ. For since, by 85, $r_1 \neq x$, it is possible to treat the pair (r_1, x) as, in 85, the pair $(r_1, \bar{x})$ was treated. The result would be to define a chain of elements $(r_1, p_1, p_2, \cdots, p_h, \cdots)$, such that $F(p_h \mid r_1 x)$, while, as before, no two elements of the chain are mutually equivalent. If any element of Σ, say v, is such that $F(v \mid r_1 x)$ and $F(v \mid r_1 \bar{x})$, we have $F(v \mid r_1 x)$ and $F(r_1 \mid vx)$, and hence $v = r_1$. See also 72. It follows that no element of the new chain except r_1 itself is equivalent to any element of the chain $(r_1, r_2, \cdots, r_n, \cdots)$.

General properties of resultants.

87. A very little consideration serves to show that a new application of principle VI, to the chains of elements now defined, will lead to still further results. Before we are prepared to consider these results, we must however survey the properties of a class of elements defined in 17. Of the infinitely numerous non-equivalent elements now known to exist in Σ, collections can be made comprising any number of elements. These may be either E-collections or O-collections. If of the former type, the collections so made can be enlarged to O-collections by the adjunction of suitable elements. To what laws are such adjunctions subject? This we are next to see.

88. If β be any collection, any element r such that $F(r \mid \beta)$, is a resultant of the collection β, by virtue of the definition stated in 17. Since r is thus any element such that $O(\beta \bar{r})$, and since the obverse of any element of β may fill the place of $\bar{r}$ in this O-collection, it is obvious that every element of β is also a resultant of β. Furthermore, if any element x forms an O-collection with any partial collection of elements of β, $\bar{x}$ is a resultant of β. For if $O(\lambda x)$, then $O(\kappa \lambda x)$, so that if (κ, λ) is the same collection as β, $F(\beta \mid \bar{x})$.

89. If $F(r|\beta)$, $F(t|\beta)$ and $F(u|rt)$ are all true there follows, by 63, $F(u|\beta)$. Hence any mediator of a pair of resultants of a given collection is itself a resultant of the collection. By the reasoning used in 85–86, it can therefore be shown that any collection containing at least two non-equivalent elements must possess an infinite number of resultants.

90. If ρ be the collection consisting of all the resultants, r_1, r_2, etc., of a given collection β, then any resultant of ρ is also a resultant of β.

Let q be such that $F(q|\rho)$. Then $O(\rho\bar{q})$, and hence $O(\bar{\rho}q)$. By the definition of a resultant, β is a collection which becomes an O-collection if any element of $\bar{\rho}$, that is, if the obverse of any one of the resultants of β, be adjoined to β. Hence β forms an O-collection in case any member of the O-collection $O(\bar{\rho}q)$, except q, is adjoined to β. By 36(5) it follows that $O(\beta\bar{q})$. Hence $F(\beta|q)$.

The collection ρ is consequently a collection which contains all of its own resultants. It also contains the resultants of all the partial collections which can be found by selecting certain elements from the collection β.

91. If y is an element of Σ, selected at pleasure, and if β is any collection, and if a determinate element c exists such that, whatever element b_n of β is chosen, c is a mediator of the pair (y, b_n), then c is also a mediator of any and every pair that can be formed by combining y with any resultant whatever, say r_x of β. For if r_x be any resultant of β, then $O(\beta\bar{r}_x)$. And if c exists such that $F(c|yb_n)$ is true, whatever element b_n of β may be chosen, then $O(\bar{c}yb_n)$ is true of every member b_n of β. Hence, by 36(5), $O(\bar{c}yr_x)$. And hence $F(c|yr_x)$.

The converse of this theorem is obvious. That is, if an element c is a mediator of *every* pair (y, r_x) consisting of y and some resultant of β, c is a mediator of (y, b_n) where b_n is any element chosen at pleasure from β. For the elements of β are themselves amongst the resultants of β.

There always exists an element c having the properties here in question, since y itself is such an element. For $F(y|yx)$ is true of every element x. But, as will soon appear, there are, in certain important classes of cases, elements possessing this property which are not equivalent to the chosen element y.

92. If γ be the collection of all those elements, c_v, which possess the property discussed in 91, viz., if γ be the collection of those elements c_v, such that $F(c_v|yr_x)$ is true for a determinate selected element y, and for every resultant, r_x, of β, then every resultant of γ is itself a member of the collection γ.

For let t be such that $F(t|\gamma)$, so that $O(\bar{\gamma}t)$. Since (y, r_x) is such that $O(yr_x\bar{c}_v)$ is true of every element $\bar{c}_v$ of $\bar{\gamma}$, while $O(\bar{\gamma}t)$, it follows that $O(yr_x\bar{t})$, by 36(5). Hence $F(t|yr_x)$.

If ρ be the collection of all the resultants of β, and if γ be the collection of all those elements, such as c_v, which with reference to a selected element y,

have the property of being, every one of them, such that $F(c_y | yr_x)$ is true of every element r_x of ρ, then, as now appears, the collections γ and ρ are collections each of which contains all of its own resultants.

Pairs of conjugate resultants.

93. We are now prepared to consider more minutely the consequences of principle VI. By that principle, if a collection ϑ possesses any complement w, so that $O(\vartheta w)$, a complement v also exists such that $O(\vartheta v)$ while, whatever element t_u of ϑ be selected, $O(vwt_u)$. Stating this requirement in terms of the F-collections, we have, as the conditions set forth in the hypothesis of the principle, the existence of an element $\bar{w}$, such that $F(\bar{w} | \vartheta)$. The consequence according to the principle is that $\bar{v}$ also exists such that $F(\bar{v} | \vartheta)$, while, since $O(vwt_u)$, $O(\bar{v}\bar{w}t_u)$, and therefore $F(t_u | \bar{v}\bar{w})$ for every element t_u of ϑ. Let $\bar{w} = q$, and $\bar{v} = r$. Then principle VI asserts that whatever resultant q of a collection ϑ be selected, there always exists a resultant of ϑ, namely r, such that every element of ϑ is a mediator (that is a resultant) of the pair (q, r). It readily follows that every resultant of ϑ is a mediator of (q, r). For if p_k be any resultant of ϑ, then $O(\vartheta\bar{p}_k)$. And since $O(vwt_u)$ is true of every element of ϑ, while $O(\vartheta\bar{p}_k)$, it follows, by 36 (5), that $O(vwp_k)$, that is $O(\bar{q}\bar{r}p_k)$, and hence $F(p_k | qr)$.

94. If q is any determinate resultant of ϑ, and if r is an element related to q in the way set forth in 93, then any element, r', such that $F(qr' | p_k)$ is true for every resultant of ϑ, i. e. any element r' such that $F(qr' | t_u)$ is true of every element of ϑ, is also such that $r' = r$. For, since r and r' are both of them resultants of ϑ, we have $F(qr | r')$, because of the definition of r; and also $F(qr' | r)$ because of the definition of r'. Hence $r' = r$ by 52 (2). We may consequently let r stand as the unique representative of the class of those resultants of ϑ which, when q is given, fulfil, with respect to q, the requirement of principle VI. With this understanding, we shall henceforth characterize r as the *conjugate resultant of q in, or with respect to* the collection ϑ. If r is given instead of q, some equivalent of q is nevertheless predetermined as a conjugate of r; and if q be selected as the unique representative of the class of elements which are equivalent to itself, we may regard the relation of q and r as wholly symmetrical; and so we may henceforth speak of the pair (q, r) as a *pair of conjugate resultants of the collection ϑ*, or more briefly as a *conjugate pair in, or with respect to, ϑ*. We shall symbolize the relation in question thus: $J(qr; \vartheta)$. This symbol is to be read as the assertion: "The pair (q, r) is a pair of conjugate resultants of the collection ϑ," or "is a conjugate pair in ϑ," or "with respect to ϑ." Were ϑ a pair, as for instance (x, y), we could write $J[qr; (x, y)]$.

95. If the collection ϑ is made to include all of the elements of Σ, the pair (q, r) becomes a pair of mutually obverse elements so that $q = \bar{r}$. This appeared already in 27, when the first use of principle VI was made. If ϑ is an O-collection, a conjugate pair (q, r) such that $O(\bar{q}\bar{r}t_n)$ while $O(\vartheta)$ is true, is a pair such, by principle II, that $O(\bar{q}\bar{r})$. Hence, in this case, $O(qr)$, and thus any pair that is a pair of conjugates with respect to an O-collection is an O-pair. If, in case of any collection ϑ, a pair (q, r), conjugate with respect to ϑ is such that $q = r$, then, since $F(t_n|qr)$, every element t_n of ϑ is such that $t_n = q = r$; and therefore, in this case, a single element may be taken as the unique representative both of the whole collection ϑ, and of all of its possible pairs of conjugate resultants (see 52 (1), and 70).

96. In case of any pair of conjugate resultants of a system β, i. e., in case of (q, r) such that $J(qr; \beta)$, the resultants of the pair (q, r), and the resultants of β, form precisely identical collections. If β is enlarged either to an O-collection, or so as to include all of the elements of Σ, the entire system Σ becomes the collection of the resultants of any one of the possible pairs of obverse elements of Σ, such, for instance as $(x, \bar{x})$; any one of these pairs being, as we now know, such that, if $O(\beta)$, $J(x\bar{x}; \beta)$.

97. If β is a given collection, and if (q, r) is a conjugate pair of its resultants, and if u is any third resultant of β, not equivalent either to q or to r, then the conjugate resultant v of u can be found by considering merely the triad (q, r, u). According to principle VI there is, namely, a resultant v of this triad* such that $F(vu|q)$ and $F(vu|r)$. Since v is a resultant of the triad, $F(qru|v)$. But meanwhile, since $J(qr; \beta)$, and $F(\beta|v)$, $F(qr|v)$ by the definition of a conjugate pair, so that from $F(qru|v)$ the element u may be stricken out. Since v, then, is such that $F(qr|v)$, while $F(vu|q)$ and $F(vu|r)$, it is easy to show that any resultant p_k of β is such that $F(p_k|vu)$. For $F(p_k|qr)$, since (q, r) is a conjugate pair. But from $F(p_k|qr)$, $F(q|uv)$ and $F(r|uv)$, follows, by 73, $F(p_k|uv)$.

If, then, a single conjugate pair of resultants of a collection β is given, viz., (q, r), the conjugate of any third resultant of β, such as u, is equivalent to the conjugate of u in the collection of the resultants of the triad (q, r, u). If β is an O-collection, or if β includes all of the elements of Σ, the conjugate of x in the triad $(x, z, \bar{z})$ is obviously $\bar{x}$.

98. If (β, x) be any collection that includes a given element, x, the conjugate resultant of x with respect to the collection (β, x) is one of the resultants of β. For let q be such a resultant. $F(q|qx)$ is in any case true of q. But q has, in addition, to be such that $F(q|\beta x)$, while, whatever element b_n, of β,

* Identical with what KEMPE calls the "unsymmetrical resultant" of the triad (q, r, u). KEMPE does not directly define our conjugate resultants in general, but builds his theory upon that of the resultants of triads.

be selected, $F(qx \mid b_n)$; i. e., $O(\bar{q}\bar{x}b_n)$. By the adjunction of all elements of β besides b_n, we obtain, hereupon, $O(\bar{q}\bar{x}\beta)$. But since $F(q \mid \beta x)$, $O(\bar{q}x\beta)$. Hence $O(\bar{q}\beta)$. Hence $F(\beta \mid q)$.

99. If β and α are such that $F(\beta \mid \alpha)$, then β and α possess at least one common resultant. For let a be an element of α, chosen at pleasure. Adjoin $\bar{a}$ to the collection β, and consider the conjugate resultant of $\bar{a}$ in the collection $(\bar{a}, \beta)$. Let this resultant be the element b. By 98, $F(\beta \mid b)$. By the definition of a conjugate resultant $F(\bar{a}b \mid b_n)$ is true of every element b_n of β. By hypothesis, however, $F(\beta \mid \alpha)$. Hence by 56, $F(\alpha \mid \bar{a}b)$. Since a is itself an element of α, the superfluous obverse $\bar{a}$ may be omitted, so that $F(\alpha \mid b)$. Hence the element b is a resultant of β and also of α.

The importance of this theorem for the geometrical application of our theory (since the theorem may be called *the theorem regarding intersections* or *transversals*) justifies a proof directly in terms of O-collections.

If, namely, $O(\delta\epsilon)$, there exists x such that $O(\delta\bar{x})$ while $O(\epsilon x)$. For let us select at pleasure any element d of δ. Consider the collections $(\bar{d}, \bar{\epsilon})$, and, with respect to that collection, define the conjugate resultant of $\bar{d}$ in $(\bar{d}, \bar{\epsilon})$. Let x be this resultant. By 98, $O(\bar{\epsilon}\bar{x})$ is true. Hence $O(\epsilon x)$. By the definition of a conjugate resultant $F(\bar{e}_r \mid x\bar{d})$ is true of every element $\bar{e}_r$ of $\bar{\epsilon}$. Hence $O(e_r x\bar{d})$ is true of every element e_r of ϵ, while $O(\delta\epsilon)$. Hence, by 24 (4), $O(\bar{\delta}x\bar{d})$. $\bar{d}$ is a repetition of some element of $\bar{\delta}$, and may be stricken out (by 37). Hence $O(\bar{\delta}x)$. Hence $O(\delta\bar{x})$, while, as above shown, $O(\epsilon x)$. So the theorem is proved.

Since this process may be repeated for every element of δ and also of ϵ, the variety of elements of the type x, in case the elements of δ and of ϵ include non-equivalent pairs, is, in general, by principle V, and 89, infinite.

The relations of pairs of conjugate resultants in various collections.

100. It is frequently important to bring the various pairs of conjugate resultants which exist in different collections into relation with one another. The procedure by which this is accomplished will lead us at once to the threshold of the ordinary algebra of logic, which, as originally developed, was based upon observing certain properties of the system Σ, in cases where this system was interpreted as a collection whose elements are either classes or propositions.

Let the conjugate resultant of y in the collection (y, β) be x, so that $J(xy; (\beta, y))$. Then, by 98, $F(x \mid \beta)$ is true. Hereupon, if we select, amongst the resultants of β, that one, say z, which is the conjugate resultant, with respect to β, of the element x just determined, so that, while $J(xy; (\beta, y))$, it is also true that $J(xz; \beta)$, then it follows that $F(yb_n \mid z)$ is true for every element b_n of β. For z is such that $F(xz \mid b_n)$, and is also such that $F(\beta \mid z)$. Since $F(b_n \mid xy)$ is true (by the foregoing) for every element b_n of β, while

$F(\beta|z)$, it follows, by 36 (5) and 57, that $F(xy|z)$; and since $F(b_a|xz)$, that is $F(b_a\bar{z}|x)$, while $F(x|z\bar{y})$, we have $F(b_a\bar{z}|z\bar{y})$, whence follows $F(z|b_a y)$.

We have, therefore, the result, that, if y be any element, and β any collection, then, in case x is such that $J(xy;\ (\beta, y))$ and z is such that $J(xz;\ \beta)$, the two elements x and z constitute a conjugate pair of the resultants of β, while just this pair stands in what we may regard as an unique relation to y. The pair (x, z) is namely such that, for the first, z is a mediator between y and any element b_a of β which may have been chosen for comparison with y and with z. By 91, z is consequently also a mediator between y and any resultant, r_c, of β, so that, whatever resultant r_e, of β, we may select $F(r_e y|z)$. Moreover, whatever resultant z' of β possesses the property just ascribed to z, must be equivalent to z. For if z' exists such that $F(z'|\beta)$, while $F(z'|r_c y)$ is true, whatever resultant r_e of β we choose to consider, then, since z itself is a resultant of β, $F(z'|zy)$, while, by the definition of z, $F(z|z'y)$, and hence $z = z'$. Therefore z may be taken as the unique representative of its own class of equivalents. Meanwhile, x possesses, with reference to β and y, the property of being a resultant of β such that every element of β is a mediator between x and y. Consequently, since $F(b_a|xy)$, $F(b_a\bar{y}|x)$, and hence x is a mediator between $\bar{y}$ and whatever element b_a of β may have been chosen. By 91, x is accordingly a mediator between $\bar{y}$ and whatever resultant, r_c, of β, may have been chosen. Whatever resultant, x', of β, possesses the property just ascribed to x, is such that $x = x'$.

101. *Conjugate limits of a collection with reference to a base.* We may sum up the result of the foregoing thus: If any element y be chosen, at our pleasure, as what we shall now call a base, and if hereupon any collection β be considered with reference to this base, then there exists a pair, and (barring for the moment the consideration of equivalent elements), a *single* pair, of conjugate resultants of β, which is so related to y and to β that (1) one of these two resultants (which we shall now symbolize by p) is such that $F(p|b_a y)$ for every element of β, and $F(p|r_c y)$ for every resultant of β; while (2) the other of these resultants, which we shall now symbolize by s, is such that $F(b_a|ys)$, and $F(r_e|ys)$, i. e., such that $F(b_a\bar{y}|s)$, and $F(r_e\bar{y}|s)$. If γ be the collection of the totality of those elements of Σ which are mediators between y on the one hand and each and every element and resultant of β, separately considered, on the other hand (see 92), then p has the property of belonging at once to the collection γ, and to the collection ρ; where ρ is, as before, the collection of all the resultants of β. Any element possessing the property of belonging at once to γ and to ρ, is equivalent to p, which may therefore be viewed, for present purposes, as the unique representative of its own class of equivalent elements. Barring equivalent elements, then, the collections γ and ρ have *only* this element p in common. If γ' be the collection of

those elements of Σ' which are mediators between $\overline{y}$ on the one hand, and each and every resultant of β, separately considered, on the other hand, then s has the property of belonging at once to the collection γ', and to the collection ρ. Any element possessing the property of belonging at once to γ' and to ρ, is equivalent to s, which may therefore be viewed, for the purposes of forming O-collections and F-collections, as the unique representative of its own class of equivalent elements. Barring equivalent elements, then, the collections γ' and ρ have *only* the element s in common. Of the pair (p, s), each resultant of β is a mediator. If, for the collection β, and for ρ, the collection of the totality of the resultants of β, the pair (p, s) alone is substituted, and if hereupon this pair is treated precisely as, in the foregoing, β itself has been treated, that is, if the resultants of (p, s) are first defined, and then their collection, viz., ρ, is compared with y, the same pair (p, s), is once more found as that pair of conjugate resultants of the collection (p, s) itself, whose relation to y is the relation heretofore characterized.

We shall now call the pair (p, s) a *pair of conjugate limits of β with reference to the base y*. For a given collection β, and for a given base y, principle VI thus requires us to define one such pair, and (barring equivalent elements), but a single pair, viz., (p, s), which may be viewed as *the* pair of conjugate limits in question. This pair, being a pair of mutually conjugate resultants of β, is symmetrically disposed with reference to the collection of the resultants of β. But, as has appeared in the foregoing, the pair (p, s) is not, in general, symmetrically disposed with respect to the enlarged collection (β, y). For p is a mediator between y and each resultant of β separately considered; while each resultant of β, separately considered, is a mediator of the pair (y, s). To mark this difference of relative position of p and s we may call: p the inferior limit of β with respect to y; s the superior limit of β with respect to y; while y is the base of this pair of conjugate limits of β. It is at once obvious that if we choose $\overline{y}$ as base instead of y, s would become the inferior, and p the superior limit of β with respect to $\overline{y}$.

102. We are now in a position to extend our result from the case of a single collection β, to the case of a set of collections β, γ, δ, etc., and in fact to the set of all possible collections of the elements of Σ. Holding a given base, y, chosen at pleasure, constant, we may consider any and all collections of the elements of Σ with reference to this one chosen base. If we do so, then, whatever collection, α, β, γ, δ, etc., we select, we shall find, by a process wholly analogous to the foregoing, that there exists an unique pair of resultants of any one such collection, such that this pair is, for that collection, the pair of conjugate limits of the collection with respect to y. Of this pair, one element is the superior, and the one the inferior limit of the collection in question, with respect to this chosen and constant base, while the choice of the base is arbitrary. The

limits are each time functions of the collection in question, and may be regarded as functions of that alone, *so long as the base is held constant.* *

If, while the base itself remains constant, a collection is altered, by the adjunction, or by the omission of elements, its limits alter or remain invariant, in ways whose laws are now to be defined. If various collections are considered with reference to the same base, their respective superior or inferior limits may be considered as new collections, with results whose laws are also to be determined. But if, instead, while any collection or sets of collections remain constant, the base is altered, so that an element y' or y'' takes the place of y, then the pairs of conjugate limits of any given collection with reference to the new base remain invariant, or alter, in accordance with still other principles (which we shall consider in chapter VI). We shall discuss these various cases in order. But in the rest of the present chapter, the base shall remain constant.

103. *Adjunction of the base to a collection.* If, to the collection β, while the base remains constant, the base y is itself adjoined, the inferior limit of (β, y) becomes equivalent to y itself; but the superior limit remains invariant. This appears from the reasoning used in 100.

104. *Adjunction of the obverse of the base; adjunction of resultants; other cases.* If the base y is held constant, and if $\bar{y}$ is thereupon adjoined to β, the inferior limit remains constant, while the superior limit of $(\beta, \bar{y})$ becomes equivalent to $\bar{y}$. The reasoning used in 100 can be employed to prove this also. — If any resultant of β is adjoined to β, or if any collection of the resultants of β is adjoined to β, the superior and inferior limits of the enlarged collection, so long as the base y is held constant, remain invariant; as appears from the reasoning used in 91, 92. Hence, by 89, if to a given collection, any mediator of any pair of its elements or of its resultants is adjoined, the limits remain constant; as they also do if the elements of a collection are repeated any multitude of times. If y is the constant base, and if any element or collection of elements of the collection γ, of 92 and 101, be adjoined to β, then, while the inferior limit, in general, alters, the superior limit of the collection remains in so far invariant. If any element or collection of elements chosen from the collection γ' of 101 be adjoined to β, this adjunction leaves the inferior limit of the enlarged collection invariant, while altering, in general, the superior limit.

Elements may be omitted from collections in a manner which is subject to these same laws. Thus, the omission of such repetitions of elements as occur in a collection does not alter either the superior or the inferior limits, etc.

* The conception that the elements usually known as the products and sums of the algebra of logic are relative to a chosen constant base (the zero of the usual algebra of logic), and that the choice of what element of Σ shall be treated as the zero-element is essentially arbitrary, is KEMPE's. But KEMPE develops this conception solely on the basis of his theory of the symmetrical and unsymmetrical resultants of *triads*. By the more general concept of the conjugate resultants of collections, I have generalized KEMPE's theory so as to be able to apply it, from the start, to collections of any multitude whatever.

The operation of determining the superior and inferior limits of a collection is obviously independent of the order in which the elements are arranged or considered, and may consequently be called a *commutative operation.*

105. *The transformation of a collection into its obverse collection.* If a collection β be transformed into its obverse collection (the base y remaining constant), if p be the inferior, and s be the superior limit of β, and if p' be the inferior and s' be the superior limit of $\bar{\beta}$ then the two equivalences hold good:

$$\bar{p} = s' \text{ ; and } \bar{s} = p' .$$

For, by the definition of an inferior limit, $F(p \mid y r_v)$ for every resultant, r_v, of β, while $F(p \mid \beta)$. Whence there follows (1) $F(\bar{p} \mid \bar{y}\bar{r}_v)$ for every resultant $\bar{r}_v$ of $\bar{\beta}$; while (2) $F(\bar{p} \mid \bar{\beta})$. From (2) it follows that $\bar{p}$ is a determinate resultant of $\bar{\beta}$. From (1) follows, by transfer, $F(\bar{p}y \mid \bar{r}_v)$, whatever resultant $\bar{r}_v$ of $\bar{\beta}$ may be chosen. Thus $\bar{p}$ is such a resultant of $\bar{\beta}$ that whatever element or resultant $\bar{r}_v$ of $\bar{\beta}$ may be chosen, this element or resultant of $\bar{\beta}$ is a mediator between $\bar{p}$ and y. Hence $\bar{p}$ is equivalent to the superior limit of $\bar{\beta}$ with respect to the base y. Hence $s' = \bar{p}$.

Furthermore, since s is the superior limit of β with respect to y, we have $F(r_v \mid sy)$, whatever resultant, r_v, of β we may choose; while, at the same time $F(s \mid \beta)$. It follows that $F(\bar{s} \mid \bar{\beta})$, so that $\bar{s}$ is a determinate resultant of $\bar{\beta}$; while $F(\bar{r}_v \mid \bar{s}\bar{y})$, i. e., $F(\bar{r}_v y \mid \bar{s})$, for every resultant $\bar{r}_v$ of $\bar{\beta}$; so that $\bar{s}$ is a mediator between y on the one hand, and each and every $\bar{r}_v$ of $\bar{\beta}$ on the other hand. Thus $\bar{s}$ is equivalent to the inferior limit of $\bar{\beta}$ for the base y. The equivalences in question, viz., $\bar{p} = s'$; and $\bar{s} = p'$, are accordingly proved.

106. *The combination of collections. The associative law in the determination of pairs of conjugate limits with respect to a constant base.* If we consider the inferior limit of a collection β_1, with respect to a base y, and also the inferior limits, with respect to the same base, of collections $\beta_2, \beta_3, \cdots, \beta_n, \cdots$, in any multitude of collections, and if π be the collection of all these inferior limits, while ω is the collection consisting of the totality of collections $(\beta_1, \beta_2, \cdots, \beta_n, \cdots)$, then the inferior limit of π is equivalent to the inferior limit of ω (the subscripts $1, 2, \cdots, n, \cdots$ are now no longer to be viewed as ordinal numbers, but merely as distinguishing marks).

The inferior limit p_ω of the collection ω is, in fact, an element such that $F(\omega \mid p_\omega)$, while if $b_n^{(k)}$ is any element of any collection β_n in the set $(\beta_1, \beta_2, \cdots, \beta_n, \cdots)$, $F(p_\omega \mid y b_n^{(k)})$. Meanwhile, if p_n is the inferior limit of any collection β_n, then $F(\beta_n \mid p_n)$, while $F(p_n \mid b_n^{(k)} y)$ for every element $b_n^{(k)}$ of β_n. If π is the collection $(p_1, p_2, \cdots, p_n, \cdots)$, then its inferior limit p_π is an element such that $F(\pi \mid p_\pi)$, while, whatever p_n may be selected, $F(p_\pi \mid p_n y)$. Since $F(p_n \mid b_n^{(k)} y)$, and $F(p_\pi \mid p_n y)$, it follows by 71, that $F(p_\pi \mid b_n^{(k)} y)$; and this latter assertion holds true, whatever element $b_n^{(k)}$ of any collection β_n

may be selected, that is, whatever element of ω may be selected. Moreover, whatever element p_n of π may be selected, since p_n is a resultant of β_n, p_n is also a resultant of ω (88). Thus $F(\omega|p_n)$, whatever element p_n of π may be selected, while $F(\pi|p_\pi)$. Hence (by 57), $F(\omega|p_\pi)$. Thus p_π is a resultant of ω, while p_π is such that $F(p_\pi|b_n^{(k)}y)$ for every element $b_n^{(k)}$ of ω. Hence $p_\pi = p_\omega$.

The operation of determining the inferior limit of a collection ω which is composed of a set of collections $(\beta_1, \beta_2, \cdots, \beta_n, \cdots)$, is consequently *associative* with respect to the operation of separately determining the inferior limit of each one of these collections $(\beta_1, \beta_2, \cdots, \beta_n, \cdots)$.

By a precisely similar reasoning, one can obviously prove that the operation of determining the superior limit, s_ω, of the collection ω, is associative, in case we separately determine the superior limits of β_1, β_2, etc., and use such determination as the basis for determining s_ω.

107. *Cross collections.* If a set of collections $(\beta_1, \beta_2, \cdots, \beta_n, \cdots)$ is given (where the multitude of collections in question is wholly unrestricted), and if a collection λ is formed by selecting, at pleasure, one element, and one only (say $b_1^{(i)}$), from β_1, one element, and one only (say $b_2^{(j)}$), from β_2, and one element, and only one, from each of the collections of the set (so that, for instance, $b_n^{(r)}$ is selected from β_n), then the collection λ_z, that is, the collection $(b_1^{(i)}, b_2^{(j)}, \cdots, b_n^{(r)}, \cdots)$ shall be called a *cross-collection* of the set $(\beta_1, \beta_2, \cdots, \beta_n, \cdots)$. As many distinct cross-collections $(\lambda_t, \cdots, \lambda_x, \cdots, \lambda_z, \cdots)$ exist as there are distinct possible combinations of elements selected one from each of the collections of the set $(\beta_1, \beta_2, \cdots, \beta_n, \cdots)$. Choose now a constant base y, to be retained throughout what follows in this and in the next section. Hereupon, let s_n be the superior limit with respect to y of the collection β_n of the original set. Let $(s_1, s_2, \cdots, s_n, \cdots)$ be the collection of all such superior limits, with respect to y, of β_1, β_2, etc. Let the collection $(s_1, s_2, \cdots, s_n, \cdots)$ be here called the collection ϑ. Let the inferior limit of ϑ, with respect to the base y, be symbolized by p_ϑ. Next, let p_z be the inferior limit of λ_z, that is, of the collection $(b_1^{(i)}, b_2^{(j)}, \cdots, b_n^{(r)}, \cdots)$. Let $(p_t, \cdots, p_x, \cdots, p_z, \cdots)$ be the collection comprising *all* such inferior limits of the possible cross-collections $(\lambda_t, \cdots, \lambda_x, \cdots, \lambda_z, \cdots)$. Let the collection $(p_t, \cdots, p_x, \cdots, p_z, \cdots)$ be called the collection ψ. Let the superior limit of ψ be symbolized by s_ψ. Then we shall next inquire how p_ϑ is related to s_ψ, that is, how the inferior limit of the collection of all the respective superior limits of a given set of collections, is related to the superior limit of the collection of all the respective inferior limits of the corresponding cross-collections. To ask this question is to inquire (in the most general form possible) whether the operation of seeking inferior limits is *distributive* with reference to the operation of seeking the superior limits of given collections.

108. *The Distributive Law.* Using the conventions of the previous section,

regarding the cross-collections, and retaining the symbols used, we are able to assert that

$$p_\vartheta = s_\psi .$$

For, by the definition of an inferior limit, whatever element p_z we may select from the collection ψ, the element p_z is such that $F(p_z | y b_n^{(r)})$ for every element $b_n^{(r)}$ which belongs to the cross-collection λ_z. And, by the rule according to which this collection λ_z has been formed, there exists, in each of the original collections β_n, an element, $b_n^{(r)}$, for which the foregoing assertion is true. Meanwhile, if we consider the relation of the element $b_n^{(r)}$ to its own collection β_n, of the original set $(\beta_1, \beta_2, \cdots, \beta_n, \cdots)$, it follows, by the definition of a superior limit, that $F(b_n^{(r)} | s_n y)$. Since, then, the element $b_n^{(r)}$ is such that $F(p_z | y b_n^{(r)})$, that is, such that $F(b_n^{(r)} | p_z \bar{y})$, while $F(b_n^{(r)} | s_n y)$, we have, by the usual rule for the elimination of an element, $F(s_n y | p_z \bar{y})$; whence follows $F(p_z | s_n y)$. Since, whatever p_z may be in question, that element of λ_z which λ_z has in common with any given collection β_n may thus be eliminated, it follows that the relation $F(p_z | s_n y)$ holds of every element p_z of the collection ψ, when the relation of p_z to any element, whatever s_n, of ϑ, is considered. Since, however, by the definition of a superior limit, $F(s_\psi | \psi)$ is true, while $F(p_z | s_n y)$ holds true of each element, p_z, of ψ, separately considered, in its relation to each element of ϑ, separately considered, it follows, by 57, that, whatever s_n be chosen, the pair (s_n, y) is such that (1) $F(s_n y | s_\psi)$.

Furthermore, since, by the definition of a superior limit, whatever p_z we may select $F(p_z | s_\psi y)$, while, by the definition of an inferior limit, $F(p_z | \lambda_z)$, it follows that, whatever λ_z we may select, $F(\lambda_z | s_\psi y)$ is always true. Since this latter relation is general, and holds for every collection λ_z, without exception, we have, for every possible cross-collection, $F(b_1^{(i)} b_2^{(j)} \cdots b_n^{(r)} \cdots | s_\psi y)$; an assertion which remains invariant and true if, instead of selecting, from any one of the collections $(\beta_1, \beta_2 \cdots \beta_n \cdots)$ the element which here appears in the collection λ_z, we select instead any other of the elements of that same collection β_1, β_2, etc. Consider, hereupon, once more, the element $b_n^{(r)}$, here selected from the collection β_n. Let $\lambda_z^{(n)}$ be what the collection λ_z becomes if that element in λ_z which is derived from the collection β_n is *omitted*. In other words, let λ_z be so subjected to partition that λ_z is the same collection as $(\lambda_z^{(n)}, b_n^{(r)})$. Then we have, from the foregoing, $F(\lambda_z^{(n)} b_n^{(r)} | s_\psi y)$, that is, by transfer, $F(b_n^{(r)} | s_\psi y \bar{\lambda}_z^{(n)})$; while this relation holds true whatever element $b_n^{(r)}$ is selected from β_n. Since, however, s_n is such that $F(s_n | \beta_n)$, while for $b_n^{(r)}$ any element whatever of β_n may be substituted, we have, by 57, the collection $(s_\psi, y, \bar{\lambda}_z^{(n)})$ such that $F(s_n | s_\psi y \bar{\lambda}_z^{(n)})$; so that, by transfer, $F(\lambda_z^{(n)} s_n | s_\psi y)$. It follows that, while, as before, $F(\lambda_z | s_\psi y)$ is true, this relation remains invariant if *we substitute for any element* $b_n^{(r)}$, *of* λ_z, *the corresponding superior limit*, s_n, *of that collection* β_n, *from which* $b_n^{(r)}$ *has*

been selected. Such substitution may be accomplished in case of any element $b_1^{(i)}$, $b_2^{(j)}$, etc., independently of whether such a substitution has been made in case of any other element of λ_z.

For since both the assertions $F(\lambda_z | s_\psi y)$, and $F(\lambda_z^{(m)} s_m | s_\psi y)$, hold good in case an element $b_m^{(q)}$ of λ_z is selected from any collection β_m which is *not* the collection β_n, we could, by a repetition of the foregoing process of reasoning, show that the relation expressed in $F(\lambda_z | s_\psi y)$ not only remains invariant whichever one of the two elements, $b_n^{(r)}$ or $b_m^{(q)}$ is selected as that element for which the corresponding s_n or s_m is substituted; but *also* remains invariant when for both of them, simultaneously, the corresponding superior limits of β_n and β_m are substituted. This result can be extended, at pleasure, therefore, to any number, or to any partial collection, or to *all* of the elements of λ_z, without regard to their multitude. The relation $F(\lambda_z | s_\psi y)$ therefore remains invariant in case, for *each and every* member of λ_z, we substitute the corresponding element s_n, viz., the superior limit of that collection β_n from which the member of λ_z which is each time in question was itself selected. Hence carrying out this substitution, we have (2) $F(\vartheta | s_\psi y)$.

But we have seen above, by (1), that $F(s_n y | s_\psi)$ is true. By adjunction it follows from this that $F(\vartheta y | s_\psi)$ is true. Hence we have at once true the *two* assertions (2) $F(y | \vartheta \bar{s}_\psi)$ and $F(y | \bar{\vartheta} s_\psi)$. Hence $F(\vartheta \bar{s}_\psi | \bar{\vartheta} s_\psi)$. Hence (3) $F(\vartheta | s_\psi)$, by 54.

By (1), therefore, $F(s_n y | s_\psi)$ is true, whatever element s_n of ϑ we may choose: while, by (3), $F(s_\psi | \vartheta)$, so that s_ψ is a resultant of ϑ. It follows that s_ψ is an element such that it is a mediator between y and every element of ϑ separately considered while s_ψ is also a resultant of ϑ. Hence, by the definition of an inferior limit, $s_\psi = p_\vartheta$; and the theorem is proved.*

109. *The second form of the distributive law.* Still retaining constant the base y, let s_z be the superior limit of any cross-collection λ_z, as such collections were defined in 107. Let the collection of all the superior limits of the cross-

* The demonstration of the distributive law here given may be regarded as a generalization of KEMPE's treatment of the symmetrical resultants of triads; although this generalization involves considerations which are somewhat peculiar to the present form of the theory of conjugate resultants. When the relations of the *F*-collections are regarded as degenerating into the specialized but more familiar relation of antecedent and consequent, the proof of the distributive law becomes subject to those difficulties whose treatment by Mr. C. S. PEIRCE, by SCHROEDER, and by Dr. HUNTINGTON, are summed up by Dr. HUNTINGTON in his *Sets of Postulates for the Algebra of Logic* (these Transactions, July, 1904). The difficulties in question are a test of the sort and of the amount of information which is surrendered when, instead of the *F*-relations viewed, so to speak, in their entirety, we confine ourselves to relations which are defined, for all the collections concerned, merely with reference to a common origin, and when we thereupon define the usual logical "sums" and "products" solely upon the basis of the antecedent-consequent relations. KEMPE's Theory, like the usual one, extends the distributive law by induction from pairs of triads to any number of cases. The present treatment, in this paper, applies the distributive law at once to collections of any multitude whatever.

collections $(s_t, \cdots, s_x, \cdots, s_z, \cdots)$, be called the collection ϕ. Let the inferior limit of ϕ be p_ϕ. Let the inferior limit of any collection β_n be p_n. Let the collection $(p_1, p_2, \cdots, p_n, \cdots)$ of all such inferior limits be the collection ϵ. Let the superior limit of ϵ be s_ϵ. Then

$$p_\phi = s_\epsilon .$$

For whatever element $b_n^{(r)}$ we may select from the cross-collection λ_z, $F(s_z y | b_n^{(r)})$ is true. And whatever collection β_n we may select, and whatever λ_z is in question, there exists in β_n an element for which the assertion $F(s_z y | b_n^{(r)})$ is true. But, in case of the collection β_n, the assertion $F(y b_n^{(r)} | p_n)$ is true. Hence, since $F(b_n^{(r)} | p_n \bar{y})$ and $F(b_n^{(r)} | s_z y)$, it follows that $F(s_z y | p_n)$; an assertion which is true of every p_n in its relation to every s_z. From $F(s_z y | p_n)$, follows $F(s_z | \bar{y} p_n)$, an assertion which is true of every element s_z of the collection ϕ, while $F(\phi | p_\phi)$. Hence the collection $(\bar{y}, p_n)$ is such that $F(\bar{y} p_n | p_\phi)$, or (1) $F(p_n | p_\phi y)$: an assertion which again holds true whatever p_n of ϵ may be selected.

Furthermore, since $F(p_\phi | y s_z)$, whatever s_z may be selected, while, by the definition of s_z, $F(s_z | \lambda_z)$, we obtain, by the usual elimination process, $F(p_\phi \bar{y} | \lambda_z)$ or $F(p_\phi \bar{y} | b_1^{(i)} b_2^{(j)} \cdots b_n^{(r)} \cdots)$. This relation remains invariant whatever element of β_1 be substituted for $b_1^{(i)}$; whatever element of β_2 be substituted for $b_2^{(j)}$; whatever element of β_n be substituted for $b_n^{(r)}$; and so on. If, as in 108, $\lambda_z^{(n)}$ be what λ_z becomes when $b_n^{(r)}$ is omitted, we have, by transfer of the entire collection $\lambda_z^{(n)}$, $F(p_\phi y \bar{\lambda}_z^{(n)} | b_n^{(r)})$; and this holds true for every element $b_n^{(r)}$ of β_n, separately considered. But $F(p_n | \beta_n)$. Hence the collection $(p_\phi, \bar{y}, \bar{\lambda}_z^{(n)})$ forms an F-collection with every element of β_n, separately considered, while $F(p_n | \beta_n)$; and hence $F(p_\phi \bar{y} \bar{\lambda}_z^{(n)} | p_n)$; that is, $F(p_\phi \bar{y} | \lambda_z^{(n)} p_n)$. And thus for any element $b_1^{(i)}$, $b_2^{(j)}$, $b_m^{(q)}$, $b_n^{(r)}$, of λ_z, there may be substituted, either simultaneously with or independently of, any of the other elements, the inferior limit, p_1, p_2, p_m, p_n of the collection $\beta_1, \beta_2, \beta_m, \beta_n$, from which the element in question is selected. If the substitution is effected simultaneously for all the elements of λ_z, it follows that $F(p_\phi \bar{y} | p_1 p_2 \cdots p_n \cdots)$; that is, by transfer (2) $F(p_\phi \bar{\epsilon} | y)$.

But by (1) $F(p_n | p_\phi y)$. By the adjunction of all the elements of ϵ besides p_n this becomes $F(\epsilon | p_\phi y)$; that is $F(\epsilon \bar{p}_\phi | y)$. By the elimination of y it follows, from (1) and (2), that $F(\epsilon \bar{p}_\phi | p_\phi \bar{\epsilon})$. Hence $F(\epsilon | p_\phi)$. The element p_ϕ is so related to ϵ that p_ϕ is a resultant of ϵ, while, by (1), whatever element p_n of ϵ be selected, $F(p_n | p_\phi y)$. Hence p_ϕ is equivalent to the superior limit of ϵ, and:

$$p_\phi = s_\epsilon ,$$

as was to be proved.

Thus the operation of seeking the superior limits is distributive with reference to the operation of seeking the inferior limits of given collections.

110. The results of 91–109 may now be summarized in the somewhat less general, but (by virtue of some of our best established mental habits) more easily apprehended form of a series of statements concerning the antecedents and consequents of the members of one or more collections.

Let an element y be chosen, and held constant, *both as the base and as the origin with reference to which antecedents and consequents* are to be determined. What we have shown is:

111. That if a given collection β be considered, then, whatever element c is an antecedent, with respect to y, of every element of β, separately considered, is also an antecedent, with respect to y, of every resultant of β (see 91). Moreover (by 101; see what is there said concerning the collection γ'), whatever element c' is a consequent of every element of β, is also a consequent of every resultant of β. If we consider: (I) the collection γ, consisting of all elements c, each of which is an antecedent of every resultant of β; (II) the collection ρ, consisting of all the resultants and of all the elements of β; and (III) the collection γ', consisting of all the elements c', each of which is a consequent of every element of β: then each of these three collections is a collection which includes all of its own resultants. Each therefore is an internally complete or "perfect" collection.

112. If β includes y, or an element equivalent to y, then the collection γ reduces to the single element y itself. If β includes $\bar{y}$, or any element equivalent to $\bar{y}$, then the collection γ' reduces to $\bar{y}$. If β is an O-collection, the same result obtains for both γ and $\bar{\gamma}$; that is γ reduces to y and γ' to $\bar{y}$ (95), all the equivalents of an element being here regarded as represented by that element.

113. The collections γ and ρ have an element in common (101). This element is what we have called p, the inferior limit of β with respect to y. Whatever element γ and ρ have in common is equivalent to p. The collections ρ and γ' have an element in common, viz., s, the superior limit of β with respect to y. Whatever element γ' and ρ have in common is equivalent to s. If all the elements of β are mutually equivalent, for instance, if they are all equivalent to b, then $b = p = s$. Otherwise, $p \neq s$. The elements p and s remain invariant in case elements of β are repeated, any multitude of times, or in case any resultants of β are adjoined to β. Moreover p remains invariant whatever elements of γ' are adjoined to β; and s remains invariant, whatever elements of γ are adjoined to β (103, 104).

114. The element ρ is definable as an element which is: (I) An antecedent of every element (and so of every resultant) of β; and (II) an element such that, whatever element of Σ is an antecedent of every element of β, is also an antecedent of p. Or again, p may be defined as that antecedent, of every resultant of β, which is also itself a resultant of β. Or finally, p may be defined as that antecedent of every element of β which is itself a consequent of every element c which agrees with p in being an antecedent of every element of β.

115. The element s is definable as an element which is: (I) A consequent of every element (and so of every resultant) of β; and (II) an element such that, whatever element of Σ is a consequent of every element of β, is also a consequent of s. Or again, s may be defined as that consequent, of every resultant of β, which is also itself a resultant of β. Or, finally, s may be defined as that consequent of every element of β which is itself an antecedent of every element c' which agrees with s in being a consequent of every element of β.

116. If β is an E-collection, no one of whose elements is equivalent to y, and no one of whose elements is equivalent to $\bar{y}$, while β itself contains at least one pair of non-equivalent elements, then, by 89, and by the reasoning of chapter V, the three collections, γ, ρ and γ', contain, each of them, an infinite number of elements.

117. Of the various collections here in question a principle holds true which is statable in general, on the basis of the foregoing, as a consequence of principle VI, and as holding throughout the system Σ: If two collections are such that one of them (say γ) includes all those elements of Σ each of which is an antecedent with respect to a given origin y, of every element of the other collection (say β), then there exists in Σ an element (the inferior limit, p, of β, with respect to y), which is at once a member of the collection γ, and also a resultant of β. Or again: If two collections α and β exist, such that every element of α is an antecedent, with respect to y, of every element of β, then there exists at least one element of Σ which is a consequent of every element of α, and which is also an antecedent of every element of β. The superior limit of α with respect to y, and the inferior limit of β with respect to y, both of them stand in this position. If they are not mutually equivalent, all the elements which are their mediators agree in possessing the character in question. Another and more restricted form of the same principle runs thus: Whatever infinite sequence κ, consisting of elements of Σ, is so definable that, with reference to a chosen origin, every element k_r of the sequence possesses a consequent k_t which also belongs to the sequence, there also exists in Σ an element which is a consequent of every element of κ. For κ is a collection of elements of Σ, and consequently possesses, by the foregoing, a superior limit, with respect to y, which also belongs to Σ. The chains of elements, defined in chapter V, consequently all of them possess superior limits belonging to Σ.*

118. It is now possible, without further difficulty, to point out that the elements of Σ possess the properties of a system of logical classes, or of entities to which the ordinary algebra of logic applies. Let the arbitrarily assumed origin

*The definite relation thus brought out between the conceptions of logical products and sums, and the conception of limits, is, so far as I know, a new feature of the present discussion. It is brought to light by defining, from the outset, these conceptions with reference to collections of unrestricted multitude.

y be taken as the 0 of the ordinary algebra of logic. Let $\bar{y}$ be taken as the 1 of that algebra (otherwise symbolized, by some, as ∞). Let the relation $p \prec_y q$ be regarded as the usual relation of logical antecedent and consequent; and let the subscript of the symbol $\prec_y$ be dropped, by virtue of that usual convention which regards the reference to 0, not as reference to an arbitrary origin, but as such that $a \prec b$ has an invariant or absolute sense. So regarded, the system Σ possesses an element, 0, such that whatever element x be chosen, $0 \prec x$; and also an element, 1, such that whatever element z be chosen, $z \prec 1$. The relation $\prec$ is transitive. If $a \prec b$ and $b \prec a$, then $a = b$. If $a \prec b$, and also $a \neq b$, the relation $\prec$ is asymmetrical. Elements such as p and s may first be viewed as determined by some given pair of elements, e. g., by the pair (a, b). The element p is then called the *product*, the element s is called the *sum* of this pair; and, in the usual symbols, one may write

$$ab = p; \qquad a + b = s.$$

The definitions of the operations of logical multiplication and of logical addition, may assume the form explained in 114 and 115. Obverses will now appear as elements each of which is what is ordinarily called the negative of the other. Since, in fact, by 95 and 112, the product of an O-collection is the origin, and its sum is the obverse of the origin (see also 80, 82), the obverse elements a and $\bar{a}$ are such that $a + \bar{a} = 1$, while $a\bar{a} = 0$. We shall have the known results (easily verifiable on the basis of the foregoing):

$$a + a = a; \qquad aa = a; \qquad a0 = 0; \qquad a + 0 = a;$$

$$a\ 1 = a; \qquad a + 1 = 1; \qquad \overline{a + b} = \bar{a}\bar{b}; \qquad \overline{ab} = \bar{a} + \bar{b};$$

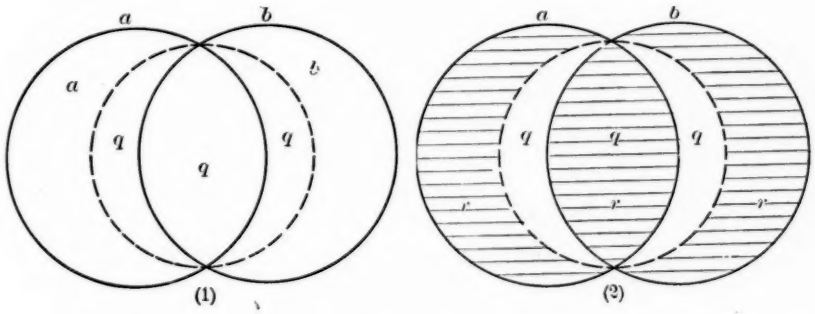
$$(ab)c = a(bc); \qquad a(b + c) = ab + ac; \qquad a + bc = (a + b)(a + c).$$

Not only are these results predetermined by the foregoing discussion, but we have in fact given to the principles in question a form much more general than the usual form by so stating the principles from the start that they apply to logical operations upon collections possessing any multitude whatever.

It follows then, that the usual algebra of logic applies without restriction to the system Σ, which is in so far identical with a totality of logical classes, whereof an infinity are mutually non-equivalent, while all are capable of an unrestricted combination by the operations of logical addition and of logical multiplication.

It is worthy of note that, in terms of the ordinary algebra, the conjugate resultants of a given collection may be defined as follows: Let the logical product of a collection β of logical elements be p. Let s be the sum of β. Then any element q such that $p \prec q \prec s$ is a resultant of β. If r is a resultant of β

such that $qr = p$, while $q + r = s$, then q and r are conjugate resultants of β. If β is the pair (a, b) represented in the diagram (1), and if q is a resultant of (a, b) so that q includes the common part of (a, b), but does not extend beyond the limits of $a + b$, then, by repeating (a, b) in the diagram (2), we may indicate, by shading, the portion of the repeated diagram where r lies, and



so the extent of r, the conjugate resultant of q, in the pair (a, b). It will be observed that q and r have the product ab in common, but supplement each other as to the remainder of $a + b$. If q expands so as to coincide with the whole of $a + b$, r shrinks to ab, and conversely.

The negatives $\bar{q}$ and $\bar{r}$ are the elements whose existence is directly asserted in principle VI.

Chapter VI. The system Σ as a generalized space-form.

119. The inquiry of the previous chapter was primarily devoted to determining what elements exist in Σ, and how they are arranged. As an incident to this research, the relations of our system Σ to the system defined in the algebra of logic was developed. But the consequences of principles V and VI, in their combination, have still other aspects. In particular, the properties of the system Σ to which we have already called attention, make its array analogous to that of the points of space. This we are next to see.

120. By definition, all of the elements of any collection are resultants of that collection. But if a collection β contains at least one pair of non-equivalent elements, and if β at the same time comprises only a finite collection of the resultants of some one of those pairs of non-equivalent elements which β contains, then there exists an infinite collection of resultants of β such that these resultants are not themselves elements of β. This appears from 89, in combination with principle V. If β is an O-collection, all the elements of Σ are resultants of β, and are also complements of β. If β is not an O-collection, Σ contains an infinity of elements which are not resultants of β. For the

obverses of the resultants of β can none of them be resultants of β; since if $F(\beta|x)$ and $F(\beta|\bar{x})$ are both of them true, $O(\beta)$. The obverses of the resultants of β are, by definition, complements of β. But if β is not an O-collection, Σ also contains an infinity of elements which are *neither* complements nor resultants of β. For let (a, b) be any pair of conjugate resultants of β, so that $J(ab; \beta)$. Since β is not an O-collection, $E(ab)$ is true (by 95). Every resultant of β is a mediator of the pair (a, b). Since $E(ab)$, $a \neq \bar{b}$ and $b \neq \bar{a}$. No element x such that $F(x|\bar{a}b)$ is true, and no element x', such that $F(x'|a\bar{b})$ is true, can be a resultant of β unless $x = b$ or $x' = a$. Since the complements of β, being obverses of the resultants of β, are mediators of the pair $(\bar{a}, \bar{b})$ we can thus define an infinity of elements which are neither complements of β nor resultants of β. All elements, x, such that $x \neq \bar{a}$, $x \neq b$ and $F(x|\bar{a}b)$; and all elements x', such that $x' \neq a$, $x' \neq \bar{b}$ and $F(x'|a\bar{b})$, are, namely, neither complements nor resultants of β. An analogous assertion holds for any other of the pairs of conjugate resultants of the collection β. Whatever pair of elements (a, b) we may choose, an infinity of pairs of elements of the form $(a, \bar{b})$, $(b, \bar{a})(c, \bar{d})$, $(d, \bar{c})(e, \bar{f})$, $(f, \bar{e})$, etc., are thereby determined, such that each pair consists of some resultant c, d, e of the pair (a, b), while the other member of each pair is the obverse of the corresponding conjugate resultant of (a, b). No mediator of any one of the pairs $(a, \bar{b})$, $(c, \bar{d})$, $(\bar{e}, f)$, etc. (except a, or c, or f, as the case may be), is a resultant of the pair (a, b).

If a, b, c, d, e, f, are mutually non-equivalent elements, while the pairs (a, b), (c, d), (e, f), are each of them pairs of conjugate resultants of the same pair, or of the same collection β, then no resultant of $(\bar{a}, b)$, or of $(a, \bar{b})$, or of $(c, \bar{d})$, or of $(\bar{c}, d)$, or of $(e, \bar{f})$, can be equivalent to any resultant of the other pairs thus defined. If x, for instance, is a mediator of $(c, \bar{d})$, so that $F(x|c\bar{d})$, $F(dx|c)$, then, with respect to x, taken as a base, the conjugate limits of (a, b), or of the collection β, in question, are, respectively c and d; c being the inferior and d the superior limit with respect to the base x, of the collection β. If therefore x' is such that $F(x'|e\bar{f})$, x' can be equivalent to x only in case each of these elements x and x', taken as base, determines, in the manner shown in 100, 101, the same pair of conjugate limits of β; in which case $c = e$, $d = f$. For if $F(dx|c)$, and $F(xf|e)$, where (c, d), and (e, f) are conjugate pairs, this result follows.

121. If we begin afresh, with a pair (a, b), and then choose a base, y, such that $F(y|a\bar{b})$, it is thus plain that, for this base y, a is the inferior, and b is the superior limit of the collection of the resultants of (a, b), and that this choice of inferior and superior limits for (a, b), *remains invariant for any base that is a mediator* of $(a, \bar{b})$, while, if the base is changed to some other pair $(c, \bar{d})$, $(d, \bar{c})$, etc. — (c, d) being a pair of conjugate resultants — the inferior

and superior limits of (a, b) (the products and sums of the ordinary algebra of logic) vary accordingly. The totality of the expressions employed in the ordinary algebra of logic to represent the relations of a system of classes, will remain invariant as to certain values, and undergoes, for other values, perfectly definite transformations, in case the base with reference to which products and sums are reckoned is altered, so that some class y takes the place which has been assigned to the zero of the ordinary algebra. If the base is, for instance, changed from some element y', such that, $F(y' | a\bar{b})$, to some element y, such that, $F(y|\bar{a}b)$, the product of the pair (a, b) is transformed into what was formerly its sum, and the sum into what was formerly the product. If the new base is an element y such that $F(y|\bar{a}\bar{b})$, the product and sum of (a, b) are transformed into a new pair of the conjugate resultants of (a, b). These transformations, somewhat analogous, for the algebra of logic, to a transformation of coördinates in a space system, lead to results which are predetermined by the F-relations of the elements of the system Σ. However the base may be changed, the product of any O-collection will be equivalent to the new base; the sum of an O-collection will be equivalent to the obverse of the base; and so on.

Meanwhile, any pair (a, b), such that $E(ab)$, while $a \neq b$, forms a means of an exhaustive classification of the elements of Σ. Given, namely, any element of Σ, say x, there is some determinate pair of resultants of (a, b), say the pair (k, l), such $F(x|k\bar{l})$. To the resultants of the pair $(k, \bar{l})$ belongs one resultant, k of (a, b), and there are also an infinite number of possible bases, for which k is product and l sum, of (a, b). No element of Σ belongs at once to two of the distinct classes thus defined by selecting pairs of conjugate resultants (i, j), (k, l), such that $i \neq k$, $i \neq l$, etc., and by then defining the class of the resultants of $(k, \bar{l})$ and of $(\bar{k}, l)$.

Or, again, one may express our present result by saying that if an element x is *not* a resultant of the pair (a, b), then there exists one and only one pair of conjugate resultants of (a, b), namely the pair (k, l), such that, if the obverse of a determinate one, say k, of these two conjugate resultants of (a, b) be chosen, $F(x|\bar{k}, l)$, while $x \neq l$.

122. If a pair of elements (a, b) be chosen such that $a \neq b$, and $E(ab)$, it is always possible to find a pair of resultants, (q, r) of (a, b), such that q and r are not mutually conjugate resultants of the pair (a, b), while $q \neq r$, and $E(qr)$, and while $F(a|qr)$ and $F(b|qr)$ are both of them false. In order to construct such a pair, it is only necessary to choose any resultant q, of the pair (a, b), such that $q \neq a$, $q \neq b$, and then to determine r such that $r \neq q$, $r \neq b$, and $F(r|qb)$. In this case, since $F(q|ab)$, and $F(r|qb)$, it is impossible that $F(a|qr)$. For if $F(a|qr)$ and $F(r|qb)$, it follows that $F(a|qb)$; while since, at the same time $F(q|ab)$, there results $q = a$, contrary to the hypothesis. Moreover, if $F(b|qr)$, and $F(r|qb)$, it follows that

$r = b$; and this again is counter to the hypothesis. Hence q and r are both of them resultants of (a, b), but neither a nor b is a resultant of (q, r).

Consider, now, the collection ρ of all the resultants of the pair (a, b), and the collection ρ_1 of the resultants of the pair (q, r). Every element of ρ_1 belongs to ρ. But there exists an infinity of elements of ρ such that no one of them either belongs to ρ_1 or is equivalent to any of the elements of ρ_1. Since the pair (q, r) is again a pair of non-equivalent elements which is not an O-pair, it is possible to determine new pairs (s, t), (u, v), (w, x), etc., without limit, such that the resultants of these pairs form a series, or chain, of collections, ρ, ρ_1, ρ_2, whereof each collection is wholly inclusive of all the elements of each later collection, while each collection contains an infinity of elements that are not included in the later collections, and that are equivalent to none of the elements so included.

It is, in the reverse direction, possible to include any collection ρ of the resultants of a given pair (so long as this is not an O-pair), in some more inclusive collection, ρ', which then may be enlarged, if necessary, to the collection of the resultants of some new pair, by considering any of the pairs of conjugate resultants of ρ'.

123. The structure of Σ is, therefore, such as to permit this endless determining of internally complete systems of resultants within systems, every such collection comprising an infinite set of elements. This being the case, the question arises whether there is also any sense in which the system Σ may be said to possess a "dimensionality" resembling that of space. The answer is that such a conception, in the system Σ, is capable of arbitrary definition in an infinite number of ways. And such a way, in fact, is suggested by the relation of any inclusive system ρ of the resultants of a pair (a, b), and any included system such as the collection ρ_1 of resultants of the pair (q, r) defined above.

Suppose, namely, that we arbitrarily define the collection of the resultants of the pair (q, r) as a one-dimensional collection, simply because the totality of these resultants is determined by the naming of the single pair of elements (q, r). In precisely the same sense, it would appear that the resultants of (a, b) or any other pair might be regarded as also of one dimension. But if we consider more carefully, it is plain that the following reason appears for a distinction between the systems ρ and ρ_1. Let m be any resultant of (q, r), such that $F(qr \mid m)$ while $m \neq q$, $m \neq r$. In ρ, that is, with respect to (a, b), m possesses a conjugate resultant n, such that $J[mn; (a, b)]$, i. e., $J(mn; \rho)$. Now it is plain that $F(qr \mid n)$ is false. For if $F(qr \mid n)$ and $F(qr \mid m)$ were both at once true, we should have, by 73, every element of ρ a mediator of (q, r), and so (q, r) would be a pair of conjugate resultants of (a, b), which is contrary to the construction as stated in 122. Consider the triad (q, r, n). Since $F(qr \mid m)$, any resultant of the pair (m, n) is a resultant of (q, r, n),

as can readily be shown by the usual elimination-process. Hence any element of ρ is a resultant of the triad (q, r, n). Hence the triad (q, r, n) possesses resultants which are not resultants of any of the pairs (q, r), (r, n), (q, n). The triad (q, r, n) resembles then a triangle, or two-dimensional complex, when viewed with reference to the pairs (q, r), (r, n), (q, n). Thus ρ can be viewed as a two-dimensional complex in relation to ρ_1. An analogous result holds whatever pair (a, q), (b, r), etc., we choose from the resultants of ρ, so long as the resultants of this selected pair form only a portion of the resultants of ρ, while elements equivalent to none of the resultants of the selected pair belong to ρ.

But we are not limited in our selection to the whole system ρ, in order to be able to define such triads. Consider next the triad (q, b, n). By 122, q and r have been so defined that $F(r|qb)$. Since, by construction, $F(m|qr)$, it follows that $F(m|qb)$. Were n also such that $F(n|qb)$, every resultant of (a, b) would be also a resultant of (q, b), which is false by construction. Hence $F(n|qb)$ is false. Were $F(q|bn)$ true, then since, as just shown, $F(m|qb)$, we should have true $F(m|bn)$, and hence, since $F(b|mn)$, it would follow that $b = m$, which is impossible by construction. For $F(m|qr)$, while $m \neq q$, and $m \neq r$. Finally, if $F(b|qn)$ were true, then, since $F(m|qb)$, it follows that $F(m|qn)$, and hence, since $F(q|mn)$, it would follow that $q = m$, which is again false by construction.

Therefore, no one of the elements of the triad (q, b, n) is a mediator of the other pair. The conjugate resultant, *in this triad*, of the element n, is an element which is a mediator of the pair (q, b) (by 98); and hence, since, m, a mediator of (q, r), is the conjugate resultant of n in ρ, it is impossible that the resultants of the triad (q, b, n) exhaust the collection ρ. Meanwhile, the triad (q, b, n) possesses resultants which are not resultants of any one of the pairs (q, b), (b, n), (q, n). And so the triad (q, b, n) may be viewed as a two-dimensional complex.

It thus follows *both* that the resultants of ρ, taken as a whole, can be viewed as the resultants of a triad, if we choose, rather than as the resultants of a pair; and that triads such as (q, b, n) can be defined, in ρ, in such wise that a triad (q, b, n) possesses resultants which are not resultants of any of its single pairs, and which are still but a part of the resultants of the system ρ. Any such triad, however, may be viewed as a two-dimensional structure.

The viewing of ρ as a two-dimensional complex with reference to ρ_1 as a one-dimensional complex, is typical of a process which can be repeated any number of times. For, since ρ_1 is itself inclusive of ρ_2, etc., ρ_1 may be viewed, with reference to these included systems, as a complex possessing two, three, or n dimensions, where n is any whole number. According as this is done, ρ comes to be viewed, with reference to a particular series of included collections, as of three, four, or $n + 1$ dimensions.

The result of the foregoing considerations is that, within any portion of the system Σ which contains at least one pair of non-equivalent elements, we can define, pairs, triads, etc., in brief, collections of any finite number of mutually non-equivalent elements, such that, if such a collection, say α, possesses n elements, there exist resultants of the whole collection which are not resultants of any partial collection of the elements, containing only $n-1$, or $n-2$, or any less number of these elements themselves.

We may call the complexes of the resultants of such collections n-dimensional complexes. But it is observable that any such complex, once given, may also be treated, by the proper choice of conjugate resultants, as a complex of the resultants of a single pair, and so as a one dimensional complex. So that all such dimensionality is entirely relative to processes and structures of the type that we have just been defining.

124. Such structures become, however, of a more positive significance if we take account of the following application.

By a *line* shall be meant a structure of the general type of the chains of 83, only completed by the insertion of certain mediators. A line shall be a collection of elements such that in case of any triad of the elements of the collection, one member of this triad is the resultant of the pair composed of the other two.

And, in particular, the lines that we are here first and mainly to consider are to be subjected to the entirely arbitrary restriction (foreign to the first principles of our system Σ, but quite capable of being satisfied by a due selection of its elements as their existence has now been established), that if any two non-equivalent entities of a line are given, *no other line*, in the set of lines that we are to consider shall at once contain both of these elements. * In other expression, let the collections which are to be called lines be so selected that, if (a, b, c) is any triad of elements belonging to the same line, $F(abc)$ is true; while, if $F(apq)$ is true and $F(bpq)$ is true, and if at the same time $F(abp)$ is false, then we shall so select that $p = q$; so that if (a, p, q) is a triad of elements belonging to one of the lines now to be selected, while (b, p, q) is a triad belonging to another of these lines, and while (a, b, p) is no linear triad at all, then we shall be required so to select that $p = q$.

125. That selections of this sort are possible the theory of the chains, as developed in 83 sqq., has already shown. Such chains as were there defined might be constructed, as we now may observe, intersecting one another any number of times. For if (c, d) be any pair of elements belonging to a chain, the resultants of (c, d) form no single chain, but lie in sets subject to principle VI, which demands the existence of conjugate resultants, not only in the collection of the resultants of the pair (c, d) itself, but in every one of the countless collections of resultants of the pairs intermediate between c and d, as

* The development is here wholly due to KEMPE's initiative.

these pairs have been characterized, in their mutual relations, in 122, 123. It is possible, within the limits of any pair of non-equivalent elements (c, d), to define any number of segments, that is of intermediate pairs (p_1, p_2), (p_2, p_3), (p_3, p_4), etc., each of which consists of mediators of (c, d), while all the elements concerned form triads such that $F(p_i p_j p_k)$ is true of any one of these triads. A chain, or rather a series of chains, can be run through such a series of intermediate pairs, according to any desired principle of selection from amongst the elements present in the various systems of resultants encountered. By virtue of the results stated in 117, a set of successive chains can be enlarged to a complete line, resembling perfectly, in its structure, a continuous geometrical line, by a mere insertion of intermediates and sets of intermediates. The special principle of selection assigned for the lines now to be considered will not only ensure that two lines have never more than one intersection, but in combination with the definition of a line will also exclude that degree of wealth of elements which forbids the arrangement of all the resultants of any pair of elements in Σ in a single linear serial order. For in the system Σ as a whole, if m and n are equivalent neither to a nor to b and are conjugate resultants of (a, b), $F(m|ab)$ and $F(n|ab)$ are both true; while $F(m|nb)$ and $F(n|mb)$ are both false. The principle laid down for the selection of line-collections will therefore forbid the inclusion in a given line of more than a single pair of conjugate resultants of any one pair. Thus, if c and d belong to the line, c and d themselves will be conjugate resultants of their own pair. And the intermediate elements of the line will be in Σ resultants of that pair. But no conjugate in (c, d) of any such resultant of (c, d) as belongs to the line, will lie in the line, except c and d themselves.

126. Since collections possessing the dimensional structure described in 123 exist in any region of Σ, it will always be possible to define systems of lines as follows: Let (d, e, f) be any triad such that it possesses resultants not contained amongst the resultants of (d, e), (e, f), (d, f). If (d, e), (e, f), (d, f), belong to lines that are amongst those lines which are here in question, and if x be any resultant of (d, e, f) which is not a resultant of (d, e), nor yet of (e, f), nor yet of (d, f), then it will be possible to regard (x, d) or (x, e), or x united with any element of the lines (d, e), etc., as constituting a new set of segments of lines. The result will be a two dimensional complex of elements. That method of construction of the n-dimensional collections or complexes of elements of Σ which has been indicated, enables us to regard these complexes, with all the lines, segments, etc., which are involved, as possessing an extent and variety of elements such as to permit us to define new sets of elements beyond any segments or bounded complexes once defined. Therefore, in selecting elements for our present purpose, we may regard these new elements as extensions of the lines and other complexes, while the dimensionality of the

complexes of lines which we may thus consider is subject altogether to our pleasure, under the conditions now in general laid down.

If, in consequence of the foregoing considerations, we compare the set of relations that we can thus define with the relations known to geometry, a natural method presents itself in the form of a juncture that may now readily be effected between our account and Dr. VEBLEN's *System of Axioms for Geometry*, (Transactions of the American Mathematical Society, July, 1904).

Dr. VEBLEN's expression "in the order ABC," becomes, in our terms, the assertion $F(b|ac)$. If we agree, in studying the constitution of our system of lines, to take explicit account only of non-equivalent elements, if we here call our elements *points*, and if we also adopt Dr. VEBLEN's definitions, his axioms appear in our statement as follows:

Dr. VEBLEN's first axiom covers our own principles III and IV, according to which our system contains a pair of elements. Axiom II of Dr. VEBLEN's set, interpreted in our terms, declares that if $F(b|ac)$, $F(b|ca)$. This needs for us, no comment. Axiom III asserts that, if $F(b|ac)$, then $F(c|ba)$ is false. Our own principles require that, in this case, $b = c$; and the axiom may therefore be regarded simply as excluding us from treating certain pairs of equivalent elements as distinct elements. This is merely a principle of selection. Axiom IV asserts that if $F(b|ac)$, then $a \neq c$. Our principles like KEMPE's, require that if $a = c$, $a = b = c$. Axiom IV, therefore, again excludes the regarding of certain equivalent elements as, for the present purpose, distinct. Axiom V of Dr. VEBLEN's set requires that if $a \neq b$, c exists such that $F(b|ac)$. This principle is provided for by our principles, which show that every pair defined in any of our sets of lines may be viewed as included in larger systems possessing linear F-relations. Axiom VI defines the important, but for us, quite arbitrary principle that governs the selection of the line-elements: "If points c, d ($c \neq d$) lie on the line (a, b), then a lies on the line (c, d)." This agrees with our foregoing statement in 124. Only, with us, this is *merely* a principle of selection. Axioms VII, IX, relate to dimensionality, and demand points existent in triads and tetrads such as we have provided for in the foregoing. For us, such requirements are permitted by the system Σ in an infinity of ways.

Axiom VIII, the "triangle transversal" axiom runs, in our terms, thus: If the triad (a, b, c) is of the two-dimensional character described in 123, and if d and e exist such that $F(e|ac)$ and $F(c|bd)$, then f exists such that $F(f|ab)$, and $F(fed)$.

This axiom for us is, if we grant a certain mode of selection of elements, a theorem, resulting from the theorem of 99 — an incidental result, as it may be called, of the theory of conjugate resultants. That is, theorem 99 secures the existence of elements which may be selected so as to verify axiom VIII.

The theorem of 99 runs that if $F(\beta|\alpha)$, β and α have at least one resultant in common.

By the hypothesis of axiom VIII, $F(c|e\bar{a})$ and $F(c|bd)$. Hence $F(e\bar{a}|bd)$. Hence follows $F(e\bar{d}|ba)$. By 99, $(e, \bar{d})$ and (b, a) have in the system Σ at least one resultant in common (they have in fact, in Σ an infinite number in common). Call this resultant f. Then f exists such that $F(f|ab)$, and $F(f|e\bar{d})$; i. e., $F(e|fd)$ so that (at least) $F(fed)$.* That the common resultants here in question should belong as points to the system of lines that we have selected from the system Σ, is itself a matter of the mode of selection used. The properties of the system Σ simply insure the possibility of such a selection.

Axiom XI, Dr. VEBLEN's form of the postulate of continuity, is provided for by our own result, holding for the system Σ in general, stated in 117. This result ensures the possibility of the continuity of the line-collections, in case we choose to select suitable sets, precisely as the same result ensures in the system Σ as a whole, the existence of "products" and "sums." Axiom X, which limits Dr. VEBLEN's system to three dimensions, is for us a perfectly possible, but again quite arbitrary limitation; and the same can be said of the parallel line axiom XII, which concerns wholly the limitation of the selection of the lines admitted into a given system.

Our own "transversal" theorem, in 99, justifies, in terms of our principles, the remark made by Mr. KEMPE, upon the basis of his postulates, to the effect that any F-collection which contains a finite number n of elements that belong to the sets selected as the lines of the foregoing discussion, represents a definite configuration of points in a space of $n-2$ dimensions.

Thus $F(abc)$ implies that (a, b, c) is a triad of points on one line. $F(ab|cd)$ is to be interpreted as follows: The pairs (a, b) and (c, d), lie by hypothesis, upon some selected pair of lines of our geometrical set. The problem is, how are these two lines to be related? The assertion $F(ab|cd)$ requires, by 99, that, in Σ, x should exist such that $F(ab|x)$ and $F(cd|x)$. If then x be viewed as one of the selected elements of the geometrical set in question, the assertion $F(ab|cd)$ may be viewed as the assertion that the lines through the segments (a, b) and (c, d) have in common a point of intersection which belongs to the mediators of (a, b) i. e., to the points of the segment (a, b), and also to the segment (c, d). On the other hand, if $F(ab|x)$ and $F(cd|x)$ are given, our principles require that $F(ab|cd)$. So that this form of assertion defines a pair of intersecting lines. The assertion $F(x|abc)$ defines a resultant of the triad (a, b, c). If this triad is to be viewed, in the way heretofore defined, as determining a complex of two dimensions, then x is a point

* KEMPE's theory is explicitly based upon two forms of the transversal theorem, assumed at the outset. For this our statement of the theory substitutes the postulated existence of conjugate pairs. What KEMPE sets at the beginning we thus reach at the end.

lying within the triangle (a, b, c). If $F(abc|def)$, wherein all the elements are mutually non-equivalent, and wherein each triad is to be viewed as a two-dimensional complex, then x exists common to these two areas, or two-dimensional complexes here in question. Thus all the intersection theorems of geometry may be stated in the form of the assertion of F-relations, with a due regard to the limitations of the classes of selected elements.

The principle of continuity is, for such a geometry, merely a principle of the selection of the elements, a principle which the system Σ permits, but does not require to be carried out.

Instead of such systems of lines as have here been selected, systems of lines whereof any two have two, three or n intersections, are perfectly permissible, so far as the system Σ is concerned. The possibility of a free, but definite variation of space-forms in a infinite number of ways, is thus provided for by the system Σ; and the outlook for a basis for generalized space-conceptions is all the more attractive, since the structure of the system Σ, based as it is upon fundamental logical principles, makes a test of the logical possibility of any proposed geometry a perfectly definite task — namely the task of seeing whether Σ actually contains complexes which are suitable to embody the desired space-form.

Since O-collections at once possess, as their resultants, all of the elements of Σ at once, no definite view of their dimensional structure is any longer possible. Hence selections suitable for space-forms must exclude O-collections; and so, as KEMPE again points out, no geometrical set contains the obverse of any of the elements of the set. It follows that "spaces," defined in the foregoing way, always occur in Σ in pairs, such that to any one space-form σ there always corresponds a space-form, or collection $\bar{\sigma}$, constituted of the obverses of the elements of σ. These two space-forms are related, in KEMPE's view, somewhat as two hemispheres.

Finally, since metrical relations can be reduced, in the known way, to ordinal relations, KEMPE has briefly pointed out (as mentioned in the introduction to this paper), that sets of the elements of Σ can be so selected that operations corresponding to the addition and multiplication of the ordinary algebra of quantity, will enable us to select elements that may be viewed (with reference to certain arbitrarily assumed constant triads of reference-elements, i. e., bases), as the sums or as the products of given pairs of elements. Hence, *without introducing new elements*, the elements of Σ, if viewed in certain ways, enable us to define, not only the algebra of logic, but the algebra of quantity.

Note on the independence of the six principles.

That principle V is independent of the other principles is proved, in 84, by the assumption of a system Σ' consisting of a single pair of obverses.

That principle IV is independent is proved by the reasoning used in 35. For if we assume a system Σ'', *all* of whose monads and possible collections are to be defined as O-collections, while the system itself comprises any arbitrarily chosen number of elements, all the principles except IV and V are satisfied by the possibly existent collections of Σ'', while principle V is satisfied "vacuously," since no pair of non-equivalent elements exist. But in such a system principle IV is false; since all the elements are, by the definition of equivalence, mutually equivalent.

Principle III, and that principle alone, would be violated by an empty system; and that principle is therefore independent.

If, instead of the O-collections, we had used, as the basis of our account of the system Σ, the indeterminate F-collections of Σ, all the principles I, III, IV, V, VI would remain true if we viewed them as statements regarding indeterminate F-collections, and could therefore have been used as principles for the system of F-collections. But principle II is false if interpreted as applying to F-collections. For $F(xx)$ is always true, since $O(x\bar{x})$. But from $F(\eta x)$ does not follow $F(\eta)$. Hence principle II is independent of the other principles.

If we consider the class of those O-collections of the system Σ which are either pairs or triads, but which contain no greater number of elements than three, we may call this class the class of the O_1-collections. For this class of collections, principle I is false, since the O_1-triads cannot be enlarged, by the adjunction of any new members. In order to apply principle II, the hypothesis of that principle must be read as applying to a collection β which, in order that it should be an O_1-collection at all, must not exceed a triad. If $O_1(\beta)$, where β is a pair or a triad (so that $O(\beta)$ is also true), and if $O_1(\delta b_n)$ is true, where b_n is any one of the two (or three) elements of β, then δ itself (by the definition of the O_1-collections), cannot exceed a pair; otherwise (δ, b_n) would be a collection of more than three elements. Principle II then becomes equivalent to the assertion that, if β is a pair or a triad, and if δ is a monad or a pair, if $O(\beta)$ is true, and if $O(\delta b_n)$ is true, then $O(\delta)$ is true. Hence principle II is true of the O_1-collections. Principles III and IV are obviously true of the elements of Σ, considered with reference to the O_1-collections. The equivalences and non-equivalences are, in fact (because of what is proved in 27–30), unchanged by the limitation of our view to the set of O_1-collections, since all equivalences and non-equivalences are already concerned in determining the relations of obverses. And principles V and VI, which require the existence of certain O-triads (such as are also O_1-collections), remain true, although the hypothesis of principle VI becomes limited, in its application, to the mention of the complements of pairs and of monads. Thus principle I is independent of the other principles.

To prove the independence of principle VI:

Consider two pairs of mutually obverse collections of elements of Σ, viz., α and $\bar{\alpha}$, such that each is a line of elements, defined as follows:

(1) Every element a_n of α is a mediator of a pair (x, y) of elements of Σ, such that neither x nor y belongs to α, while $E(xy)$ is true. Thus $F(a_n|xy)$ is true of every element a_n of α.

(2) Whatever triad of elements of α be chosen, say the triad (a_m, a_n, a_r), $F(a_m a_n a_r)$ is true of this triad.

(3) Whatever element a_n of α is chosen, a_q and a_r exist, belonging to the collection α, and such that $F(a_q|xa_n)$ and $F(a_r|ya_n)$, while $a_q \neq a_n$, and $a_r \neq a_n$.

(4) No two distinct elements of α are mutually equivalent elements of Σ.

(5) Whatever pair (a_q, a_r) be chosen from amongst the elements of α, a_m exists such that $F(a_m|a_q a_r)$ is true.

From this definition of the line α, the properties of the obverse collection $\bar{\alpha}$ at once follow. The elements of $\bar{\alpha}$ are mediators of the pair $(\bar{x}, \bar{y})$. No element of $\bar{\alpha}$ can be equivalent to any element of α; for if any mediator of $(\bar{x}, y)$ is equivalent to a mediator of (x, y), then $O(xy)$ is true, which opposes condition (1). If then c and d be distinct elements, chosen in any way from the total collection $(\alpha, \bar{\alpha})$, $c \neq d$ by construction.

Whatever pair (c, d) of elements of $(\alpha, \bar{\alpha})$ be chosen, it follows that there exists g, also belonging to $(\alpha, \bar{\alpha})$, and such that $g \neq c$, $g \neq d$, and $F(g|cd)$. For if c and d both belong to α, the existence of g, as an element of α, follows directly from condition (5). If c and d be both chosen from $\bar{\alpha}$, a precisely analogous result holds true. But if c be chosen at random from α, and d from $\bar{\alpha}$, then let $c = a_n$, and let $d = \bar{a}_q$, where $\bar{a}_q$ is such that its obverse a_q is such that $F(a_q|xa_n)$. Hereupon, choose in α an element a_t such that $F(a_t|a_q x)$. This, by condition (3), is always possible. Since $F(a_q|xa_n)$ and $F(a_t|xa_q)$, it follows that $F(a_t|a_q \bar{a}_n)$ is true; hence $F(\bar{a}_t|\bar{a}_q a_n)$ is true; and thus, if $g = \bar{a}_t$, $F(g|cd)$ is true of an element g which belongs to $(\alpha, \bar{\alpha})$. Similarly, were it true that $c = a_n$, and $d = \bar{a}_r$, where $\bar{a}_r$ is an element such that $F(a_r|ya_n)$, then we might choose a_w such that $F(a_w|ya_r)$. This, by condition (3), is also possible. Eliminating y we have $F(a_w|a_r \bar{a}_n)$; i. e., $F(\bar{a}_w|\bar{a}_r a_n)$; so that, if $g = \bar{a}_w$, we again have g belonging to the total collection $(\alpha, \bar{\alpha})$, and such that $F(g|cd)$ is true. Since all pairs of elements (c, d) thus chosen from $(\alpha, \bar{\alpha})$ are, by construction pairs of mutually non-equivalent elements; since g always exists, belonging to $(\alpha, \bar{\alpha})$ and such that $F(g|cd)$ is true; and since $\bar{g}$ also belongs to $(\alpha, \bar{\alpha})$; it follows that the collection $(\alpha, \bar{\alpha})$ contains sufficient elements to satisfy, with respect to any pair of elements of $(\alpha, \bar{\alpha})$, the demands of principle V, without going beyond the elements of this collection $(\alpha, \bar{\alpha})$ itself. That is, whatever pair of elements of $(\alpha, \bar{\alpha})$ be chosen, an element of $(\bar{\alpha}, \alpha)$ exists which is no obverse of either of the elements of the pair, and which forms an O-triad when adjoined to the pair.

Hereupon let us consider a system Σ_α, which shall consist solely of the elements of the collection $(\alpha, \bar{\alpha})$. Let there be formed, of the elements of this system Σ_α, collections which we shall call the O_α-collections. These O_α-collections shall be identical with those collections of the elements of $(\alpha, \bar{\alpha})$ which are O-collections in Σ. The O_α-collections of the system Σ_α will now conform, by construction, to principles I–V of the system Σ.

But, in case of the O_α-collections, principle VI will be violated. For (returning to the system Σ) consider any element a_n of α; and consider the totality of those elements a_q, belonging to α, and such that $F(a_q | x a_n)$ is true in respect of the system Σ. Call this totality the collection α'; so that α' is the collection of those elements of α which, in Σ, are mediators between a_n and x. By condition (1), x itself does not belong to α, and so does not belong to α'. Now, in the system Σ, the element a_n, since it belongs, by construction, to the collection α', possesses a conjugate resultant with respect to α'. Let r be this resultant, so that, in Σ, $\mathcal{M}_r(a_n r; \alpha')$ is true. It is manifest that r is an element such that $F(r | a_n x)$. Yet r is no element of the collection α' or of the total collection $(\alpha, \bar{\alpha})$. For if r were an element of α, then, by conditions (3) and (4), a_k would exist, belonging to α, and such that $F(a_k | rx)$, while $a_k \neq r$, and $a_k \neq x$. But a_k would also belong to α', and by the definition of a conjugate resultant $F(a_k | r a_n)$ would be true. Now $F(r | a_n x)$ is true; and from $F(a_k | rx)$ and $F(r | a_n x)$ follows $F(r | a_k a_n)$. If, at the same time, $F(a_k | r a_n)$, $a_k = r$; which is contrary to the hypothesis. Thus r is no element of α, and also cannot be any element of $\bar{\alpha}$. For since $F(r | a_n x)$, if there existed $\bar{a}_q$ such that $r = \bar{a}_q$, $F(\bar{a}_q | a_n x)$, and hence $F(\bar{a}_q | xy)$, would be true of some element a_q of which $F(a_q | xy)$ would also be true. In that case $O(xy)$, which is contrary to condition (1).

Since r does not belong to α nor yet to $\bar{\alpha}$, r does not exist in the system Σ_α. There is then, in Σ_α, no element capable of meeting the requirements of principle VI as applied, in this system Σ_α, to the partial collection α'. So principle VI fails in Σ_α; and is therefore independent of the other principles.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

ON MULTIPLE INTEGRALS

BY

JAMES PIERPONT

§1. *Introduction.*

In the theory of multiple proper integrals the chief difficulty arises in the consideration of the frontier points of the field of integration $\mathfrak{A}$. The older writers as GAUSS, DIRICHLET, JACOBI and CAUCHY, supposed the field $\mathfrak{A}$ to be bounded by a finite number of surfaces of simple nature. In later years the theory of point aggregates has made it possible to consider more general classes of fields. STOLZ, PRINGSHEIM † and JORDAN ‡ have made valuable contributions in this direction. The work of JORDAN is particularly important as he has not only blocked out the way for multiple integrals in any number of variables but has treated the case of two variables with great generality and completeness.

Two fundamental theorems in this theory relate to the reduction of a multiple integral to an iterated integral and the change of variables in a multiple integral. As far as the writer was aware a year ago when he began to consider the general treatment of these questions, none of the existing methods admitted an extension to n variables, at least not without considerable complication. It occurred to him that by properly generalizing the definition of an integral the grave difficulties which the frontier points create might very readily be obviated.

The theory of integration which is here given, was presented by the author at the April meeting of the American Mathematical Society in 1904. At that time Professor MOORE called his attention to a paper of STOLZ's, § published in 1897, where not only the integrals considered in the present paper but also another class of integrals are introduced. The importance of the integrals first mentioned seems entirely to have escaped STOLZ's observation. At any rate he seems to have made no further use of them, and to have considered them more as curiosities. The only theorem he establishes is theorem 2 of the present paper.

* Presented to the Society, April 30, 1904. Received for publication May 19, 1905.

† Sitzungsberichte der Münchener Akademie, vol. 29 (1899), p. 39.

‡ Journal de Mathematiques, ser. 4, vol. 8 (1892), p. 69, and Cours d'Analyse, vol. 1.

§ *Zwei Grenzwerthe, von welchen das obere Integral ein besonderer Fall ist*, Sitzungsberichte der Wiener Akademie, vol. 106, abt. II, *a*, p. 453.

§ 2. *Preliminary definitions.*

It is convenient to represent the points $x = (x_1, \cdots, x_m)$ of an m-way space R_m by a set of points marked off on m axes. Let us divide each of these axes into intervals of length $\leqq d$; in such a manner, however, that no finite segment of any axis contains more than a finite number of intervals. We shall say that we have effected *a rectangular division of space of norm* d. It divides R_m into rectangular *cells* whose sides are of length $\leqq d$. The boundaries of the cells are planes parallel to the coördinate-planes. When the lengths of the intervals are all the same, these cells become cubes. The corresponding division may be called *cubical*.

Let $\mathfrak{A}$ be a limited point aggregate, i. e., let no coördinate of any point of $\mathfrak{A}$ be greater numerically than some fixed number.

Let $f(x_1, \cdots, x_m)$ be a limited one-valued function defined over $\mathfrak{A}$. Let D be a rectangular division of space of norm d. Let $d_1, d_2, d_3, \cdots$ denote those cells of D containing at least one point of $\mathfrak{A}$. Without ambiguity we may denote the volume of these cells by the same letters.

Let

$$M_i = \max_{d_i} f, \qquad m_i = \min_{d_i} f, \qquad \text{in } d_i;$$

and

$$|f(x_1, \cdots, x_m)| \leqq F, \qquad \text{in } \mathfrak{A}.$$

We shall set

$$\bar{S}_D = \sum_i M_i d_i, \qquad \underline{S}_D = \sum_i m_i d_i, \tag{1}$$

which may be called the *upper* and *lower sums* for the division D. It is sometimes necessary to consider such sums over different aggregates as $\mathfrak{A}, \mathfrak{B}, \cdots$. These may be indicated by

$$\bar{S}_{\mathfrak{A}D}, \bar{S}_{\mathfrak{B}D}, \cdots,$$

where the subscripts indicate that the sums (1) are taken over $\mathfrak{A}, \mathfrak{B}, \cdots$ respectively.

The difference $M_i - m_i$ is the oscillation of f in the cell d_i. We denote it by

$$\operatorname{osc} f_i.$$

We remark here once for all that all functions and point aggregates considered in this paper are supposed to be limited. The case when either or both are unlimited will be considered in a later paper.

Let $a = (a_1, \cdots, a_m)$, $b = (b_1, \cdots, b_m)$ be two points. The distance between a and b is defined by

$$\operatorname{dist}(a, b) = \sqrt{(a_1 - b_1)^2 + \cdots + (a_m - b_m)^2}.$$

If c is any other point, we have

$$\text{dist}\,(a, c) \leqq \text{dist}\,(a, b) + \text{dist}\,(b, c).$$

Let $\mathfrak{B}$ be a partial aggregate of $\mathfrak{A}$. If the distance between any frontier point $\mathfrak{A}$ and any frontier point of $\mathfrak{B}$ is $\geqq \rho > 0$, we say $\mathfrak{B}$ is an *inner aggregate* of $\mathfrak{A}$; also $\mathfrak{A}$ is an *outer aggregate* of $\mathfrak{B}$.

If $\mathfrak{A}$ contains all its limiting points, it is *complete*. If it is complete and contains only limiting points, it is *perfect*.

A point aggregate embracing none of its frontier points is called a *region*.

The points x such that dist $(a, x) \leqq \delta$ form the *domain of a of norm δ*. We denote it by $D_\delta(a)$.

§ 3. *Fundamental theorems.*

THEOREM 1. *Let $f(x_1, \cdots, x_m) \geqq 0$ be defined over the field $\mathfrak{A}$. Let*

$$\bar{S} = \min S_D$$

with respect to all rectangular division D of norm d, however small. Then

$$\lim_{d=0} \bar{S}_D = \bar{S}. \tag{1}$$

The points of $\mathfrak{A}$ lie in a certain cube $\mathfrak{C}$ of edge C. The representation of $\mathfrak{C}$ is formed of m segments $\mathfrak{C}_1, \cdots, \mathfrak{C}_m$ on the $x_1, \cdots, x_m$ axes. We may suppose $\mathfrak{C}$ taken so large that no point of $\mathfrak{A}$ comes within a certain distance of any side of $\mathfrak{C}$.

Since $\bar{S}$ is obviously finite, there exists for each $\epsilon > 0$ a division Δ such that

$$\bar{S} \leqq \bar{S}_\Delta < \bar{S} + \frac{\epsilon}{2}. \tag{2}$$

Let D be an arbitrary division of norm d. Let us superimpose Δ on D, forming a division E. The division E is formed by interpolating certain points, let us say at most μ in each of the segments $\mathfrak{C}_1, \cdots, \mathfrak{C}_m$. The interpolation of one of these points may be interpreted as passing a plane parallel to one of the sides of $\mathfrak{C}$. Its effect is to subdivide certain of the cells of $\mathfrak{C}$. The volume of the cells so effected is

$$\leqq d C^{m-1}.$$

Hence the superimposition of Δ on D, being equivalent to passing at most $m\mu$ planes, affects cells of $\mathfrak{C}$ belonging to the original division D, whose volume is

$$V < m\mu d C^{m-1}. \tag{3}$$

Let Δ subdivide d_i, a cell of D containing points of $\mathfrak{A}$, into the cells

$$d_{i1},\ d_{i2},\ \cdots$$

containing points of $\mathfrak{A}$ and into cells

$$\delta_{i1},\ \delta_{i2},\ \cdots$$

containing no points of $\mathfrak{A}$. Then

$$\bar{S}_D = \sum_i M_i d_i + R; \qquad \bar{S}_E = \sum_{ik} M_{ik} d_{ik} + R,$$

where R denotes the sum of those terms common to $\bar{S}_D$ and $\bar{S}_E$ corresponding to cells of D unaffected by the division Δ. But

$$d_i = \sum_{ik} d_{ik} + \sum_{ik} \delta_{ik};$$

hence

$$\bar{S}_D = \sum M_i d_{ik} + \sum M_i \delta_{ik} + R.$$

Thus

$$0 \leqq \bar{S}_D - \bar{S}_E \leqq \sum (M_i - M_{ik}) d_{ik} + \sum M_i \delta_{ik},$$

$$< F \sum d_{ik} + F \sum \delta_{ik} = FV,$$

$$< m\mu d F C^{m-1},$$

by (3). If we take

$$d' < \frac{\epsilon}{2m\mu F C^{m-1}},$$

we have

$$\bar{S}_D < \bar{S}_E + \frac{\epsilon}{2},$$

for any $d \leqq d'$. But regarding E as formed by superimposing D on Δ, we have

$$\bar{S}_E \leqq \bar{S}_\Delta$$

since $f \geqq 0$.

Thus

$$\bar{S} \leqq \bar{S}_D < \bar{S}_\Delta + \frac{\epsilon}{2} < \bar{S} + \epsilon,$$

which proves (1).

In the same way we establish:

Let $f(x_1 \cdots x_m) \leqq 0$. *Let*

$$\underline{S} = \max \underline{S}_D$$

with respect to all rectangular divisions D. *Then*

$$\lim_{d=0} \underline{S}_D = \underline{S}.$$

THEOREM 2. *Let* $f(x_1, \cdots, x_m)$ *be defined over the field* $\mathfrak{A}$. *Then the limits*

$$\lim_{d=0} \underline{S}_D, \qquad \lim_{d=0} \bar{S}_D$$

exist and are finite. They are called respectively the lower and upper integrals of f over $\mathfrak{A}$ and are denoted by

$$\underline{\int}_{\mathfrak{A}} f d\mathfrak{A} = \underline{\int}_{\mathfrak{A}} f dx_1 \cdots dx_m; \qquad \bar{\int}_{\mathfrak{A}} f d\mathfrak{A} = \bar{\int}_{\mathfrak{A}} f dx_1 \cdots dx_m.$$

This is a corollary of theorem 1. We have only to introduce an auxiliary function

$$g(x_1, \cdots, x_m) = f(x_1, \cdots, x_m) + c$$

and take the constant c so that g is either $\leqq 0$ or $\geqq 0$ in $\mathfrak{A}$ according as we wish to show the existence of the first or second of the above limits.

When

$$\underline{\int}_{\mathfrak{A}} f d\mathfrak{A} = \bar{\int}_{\mathfrak{A}} f d\mathfrak{A},$$

we say f is integrable in $\mathfrak{A}$; the common value of these two limits we denote by

$$\int_{\mathfrak{A}} f d\mathfrak{A} = \int_{\mathfrak{A}} f dx_1 \cdots dx_m$$

and call it the integral of f over $\mathfrak{A}$.

We may now prove at once the following theorem:

THEOREM 3. *Let* D *be any rectangular division of norm* d, *and* ξ_i *any point of* $\mathfrak{A}$ *in the cell* d_i. *Then if* f *is integrable in* $\mathfrak{A}$,

$$\lim_{d=0} \sum f(\xi_i) d_i = \int_{\mathfrak{A}} f d\mathfrak{A},$$

however the D*'s and* ξ*'s are chosen. Conversely if this limit exists,* f *is integrable in* $\mathfrak{A}$.

It is now easy to show that the usual criteria for integrability hold for the present case.

An important class of integrable functions are those having limited variation. In fact, let $f(x_1, \cdots, x_m)$ be defined over $\mathfrak{A}$. Let D be a cubical division of norm $d \leqq d_0$. If there exists a positive number ω such that

$$\sum_i d^{m-1} \operatorname{osc} f_i < \omega,$$

however D be chosen, we say f has *limited variation* in $\mathfrak{A}$; otherwise f has *unlimited variation.*

THEOREM 4. *If $f(x_1, \cdots, x_m)$ has limited variation in $\mathfrak{A}$, it is integrable in $\mathfrak{A}$.*

For,

$$\sum d_i \operatorname{osc} f_i \leqq \sum d^m \operatorname{osc} f_i < \omega d .$$

But

$$\lim_{d=0} \omega d = 0 .$$

Hence f is integrable.

§ 4. *Content of point aggregates.*

In all the following work the notion of *content* is fundamental. This has been defined in various ways, e. g., by CANTOR, BOREL, JORDAN. The definition we adopt is equivalent to that of CANTOR and JORDAN, although on the face, it is different. We have chosen it because of its relation to the definition of a multiple integral just given.

Let $\mathfrak{A}$ be any limited point aggregate in R_m. Let us effect a division D of space of norm δ. Let

$$d_1, \ d_2, \ d_3, \ \cdots$$

denote those cells of D containing at least one point of $\mathfrak{A}$, while

$$d_1', \ d_2', \ d_3', \ \cdots$$

denote those cells *all* of whose points lie in $\mathfrak{A}$.

Then the limits

$$\bar{\mathfrak{A}} = \lim_{\delta=0} \sum d_i ; \qquad \underline{\mathfrak{A}} = \lim_{\delta=0} \sum d_i'$$

exist and are finite.

For, let

$$g(x_1, \cdots, x_m) = 1 \text{ at points of } \mathfrak{A},$$

$$= 0 \text{ at other points.}$$

Then obviously

$$\bar{\mathfrak{A}} = \overline{\operatorname{cont}}\, \mathfrak{A} = \bar{\int_{\mathfrak{A}}} g d\mathfrak{A} ; \qquad \underline{\mathfrak{A}} = \underline{\operatorname{cont}}\, \mathfrak{A} = \underline{\int_{\mathfrak{A}}} g d\mathfrak{A} .$$

The number $\bar{\mathfrak{A}}$, $\underline{\mathfrak{A}}$ are called *the upper and lower content of* $\mathfrak{A}$. When

$$\bar{\mathfrak{A}} = \underline{\mathfrak{A}},$$

their common value is called the *content of* $\mathfrak{A}$; in this case we say $\mathfrak{A}$ is *measurable.*

We observe that criteria for the integrability of $f(x_1, \cdots, x_m)$ give at once criteria for the measurability of $\mathfrak{A}$. When no ambiguity can arise, we may denote the content of $\mathfrak{A}$ by the same letter, otherwise we denote it by

$$\text{cont } \mathfrak{A}.$$

When cont $\mathfrak{A} = 0$, we say that $\mathfrak{A}$ is *discrete*. We show at once:

For $\mathfrak{A}$ to be measurable, it is necessary that the content of its frontier points $\mathfrak{F}$ be 0; i. e that $\mathfrak{F}$ be discrete.

With the points of an aggregate $\mathfrak{A}$ let us form partial aggregates $\mathfrak{A}_1, \cdots, \mathfrak{A}_s$, such that the aggregate of the common frontier points of any two of these aggregates is discrete. We shall say that we have divided $\mathfrak{A}$ into the *unmixed aggregates* $\mathfrak{A}_1, \cdots, \mathfrak{A}_s$.

THEOREM 5. *If we divide $\mathfrak{A}$ into the unmixed aggregates $\mathfrak{A}_1, \cdots, \mathfrak{A}_s$, then*

$$\bar{\mathfrak{A}} = \bar{\mathfrak{A}}_1 + \cdots + \bar{\mathfrak{A}}_s; \qquad \underline{\mathfrak{A}} = \underline{\mathfrak{A}}_1 + \cdots + \underline{\mathfrak{A}}_s.$$

For, let D be a rectangular division of norm δ. Let $\bar{\mathfrak{F}}_D$ be the volume of those cells of D which contain points of more than one of the aggregates $\mathfrak{A}_1, \cdots, \mathfrak{A}_s$. Let $\bar{\mathfrak{A}}_{i,D}$ be the volume of those cells containing points of $\mathfrak{A}_i$ $(i = 1, 2, \cdots, s)$. Then

$$\bar{\mathfrak{A}}_D \leqq \bar{\mathfrak{A}}_{1,D} + \cdots + \bar{\mathfrak{A}}_{s,D} \leqq \bar{\mathfrak{A}}_D + s\bar{\mathfrak{F}}_D. \tag{1}$$

By hypothesis

$$\lim_{\delta=0} \bar{\mathfrak{F}}_D = 0.$$

Hence passing to the limit in (1), we get the first half of the theorem. The other half follows similarly.

In showing that certain aggregates have content, the two following theorems are sometimes useful. They rest on the notion of a total difference quotient.

Let $f(x_1, \cdots, x_m)$ be defined over $\mathfrak{A}$. Let $x = (x_1, \cdots, x_m)$ be an arbitrary but fixed point of $\mathfrak{A}$ while $x' = (x_1 + h_1, \cdots, x_m + h_m)$ is a variable point of $\mathfrak{A}$. Let $\Delta f = f(x') - f(x)$, and $\Delta x = \text{dist}\,(x, x')$. The quotient

$$\frac{\Delta f}{\Delta x}$$

may be called the *total difference quotient of f at x*. The point x' may, or may not, be restricted to remain near x, as required.

Let

$$y_1 = f_1(x_1, \cdots, x_m), \quad \cdots, \quad y_n = f(x_1, \cdots, x_m)$$

be defined over $\mathfrak{A}$. As $x = (x_1, \cdots, x_m)$ ranges over $\mathfrak{A}$, the point $y = (y_1, \cdots, y_n)$

will range over an aggregate $\mathfrak{B}$ in an n-way space R_n. We call $\mathfrak{B}$ the *image of* $\mathfrak{A}$.

THEOREM 6. *If* $\mathfrak{A}$ *is discrete, and the functions*

$$y_1 = f(x_1, \cdots, x_m), \cdots, y_n = f_n(x_1, \cdots, x_m) \qquad (n = m + p \geqq m)$$

have limited total difference quotients, the image $\mathfrak{B}$ *of* $\mathfrak{A}$ *is discrete.*

THEOREM 7. *Let the functions*

$$y_1 = f_1(x_1, \cdots, x_m), \cdots, y_n = f_n(x_1, \cdots, x_m) \qquad (n = m + r > m)$$

have limited total difference quotients in $\mathfrak{A}$ *except at points of a discrete aggregate* Δ. *In the cells of any cubical division of norm* $d \leqq d_0$, *let at least* m *of these difference quotients remain limited. Then* $\mathfrak{B}$, *the image of* $\mathfrak{A}$, *is discrete.*

We prove only theorem 6, as the demonstration of 7 is similar. Let us effect a cubical division D of the x-space of norm d. Since the difference quotients are limited,

$$|\Delta f_i| < dG \qquad (i = 1, 2, \cdots, m),$$

when x ranges over any one of the cells d_i of D. Hence each coördinate remains in an interval of length $< dG$ as x ranges over the points of $\mathfrak{A}$ in d_i. Hence $y = (y_1, \cdots, y_n)$ remains within a cube of volume $d^n G^n$, and thus

$$\bar{\mathfrak{B}} < \sum_{\mathfrak{A}} d^n G^n = d^p G^n \sum_{\mathfrak{A}} d^m = d^p G^n \bar{\mathfrak{A}}_D.$$

As $\lim \bar{\mathfrak{A}}_D = 0$, it follows that $\bar{\mathfrak{B}} = 0$; hence cont $\mathfrak{B} = 0$.

It is convenient to extend the terms cells, division of space into cells, etc., as follows. Let us suppose the points of any aggregate $\mathfrak{A}$, which may be R_m, arranged in partial aggregates which we shall call *cells* and which have the following properties:

1°. There are only a finite number of cells in any limited portion of space.

2°. The frontier of each cell is discrete.

3°. Each cell lies in a cube of side $\leqq \delta$.

We shall call this a *division of* $\mathfrak{A}$ *of norm* δ.

Let Δ be such a division of space. As in the case of rectangular division, $\bar{\mathfrak{A}}_\Delta$ may denote the contents of all the cells which contain at least one point of $\mathfrak{A}$; while $\underline{\mathfrak{A}}_\Delta$ may denote the content of those cells *all* of whose points lie in $\mathfrak{A}$.

It is now easy to show that

THEOREM 8. $\lim_{\delta=0} \bar{\mathfrak{A}}_\Delta = \bar{\mathfrak{A}}, \qquad \lim_{\delta=0} \underline{\mathfrak{A}}_\Delta = \underline{\mathfrak{A}}.$

If $\mathfrak{A}$ does not contain all of its limiting points, these may be adjoined to $\mathfrak{A}$. The resulting aggregate is called the *completed aggregate.*

THEOREM 9. *An aggregate* $\mathfrak{A}$ *and its completed aggregate* $\mathfrak{B}$ *have the same upper content.*

For, let us effect a rectangular division D of norm d. The cells containing points of $\mathfrak{B}$ fall into two classes: 1° those cells $d_1, d_2, \cdots$ containing points of $\mathfrak{A}$; 2° those cells $e_1, e_2, \cdots$ containing no point of $\mathfrak{A}$. Each of these latter cells as e_i is contiguous to at least one cell d_k. If $e_i, \cdots$ are contiguous to d_k we will join them to d_k to form a new cell δ_k, in such a way that each e-cell has been joined to some one d-cell. The cells $\delta_1, \delta_2, \cdots$ together with the cells $d_1, d_2, \cdots$ which remain unchanged by this process of consolidation define a division Δ. The norm δ of this division converges to 0 with d.

But for the division Δ,

$$\bar{\mathfrak{A}}_\Delta = \bar{\mathfrak{B}}_\Delta;$$

and by theorem 8,

$$\lim_{\delta=0} \bar{\mathfrak{A}}_\Delta = \bar{\mathfrak{A}}.$$

Hence

$$\bar{\mathfrak{B}} = \bar{\mathfrak{A}}.$$

We may now show at once:

THEOREM 10. *If* $\mathfrak{A}$ *is measurable, its completed aggregate is also measurable; and both have the same content.*

THEOREM 11. *If* $\mathfrak{A}$ *is measurable, it has the same content as its derivative* $\mathfrak{A}'$.

Connected with any complete aggregate $\mathfrak{A}$ of upper content $\bar{\mathfrak{A}} > 0$, is an aggregate $\mathfrak{B}$, obtained from $\mathfrak{A}$ by a process of sifting, and therefore it may be called the *sifted aggregate of* $\mathfrak{A}$.

Let $D_1, D_2, \cdots$ be a set of rectangular divisions of space each formed by superimposing a division on the preceding. Let the norms of these divisions converge to zero. The division D_1 effects a division of $\mathfrak{A}$ into unmixed partial aggregates. Let $\mathfrak{A}_1$ denote those partial aggregates whose upper content is > 0. Then $\bar{\mathfrak{A}}_1 = \mathfrak{A}$. Similarly D_2 defines a partial aggregate of $\mathfrak{A}_1$, such that $\bar{\mathfrak{A}}_2 = \mathfrak{A}$. As this process goes on, the cells of D_n containing points of $\mathfrak{A}_n$ diminish in size and, for $n = \infty$, converge to a set of points $\mathfrak{B}$. The upper content of the points of $\mathfrak{A}$ in the domain of any point of $\mathfrak{B}$ is > 0. Thus each point of $\mathfrak{B}$ is a limiting point of $\mathfrak{A}$, and hence is a point of $\mathfrak{A}$. It is now easy to show that $\mathfrak{B}$ contains no isolated points, and is hence perfect. Moreover

$$\bar{\mathfrak{B}} = \bar{\mathfrak{A}}.$$

§ 5. *Some properties of multiple integrals.*

THEOREM 12. *Let* $f(x_1, \cdots, x_m)$ *be defined over the field* $\mathfrak{A}$. *Let* Δ *be any division of norm* δ *into cells* $\delta_1, \delta_2, \cdots$, *not necessarily rectangular. Let* $\mathfrak{M}_i, \mathfrak{m}_i$ *be respectively the maximum and minimum of* f *in* δ. *Then*

$$\lim_{\delta=0} \bar{S}_\Delta = \lim \sum \mathfrak{M}_i \delta_i = \int_{\mathfrak{A}}^{\bar{}} f d\mathfrak{A}. \tag{1}$$

$$\lim_{\delta=0} \underline{S}_\Delta = \lim \sum \mathfrak{m}_i \delta_i = \int_{\underline{\mathfrak{A}}} f d\mathfrak{A}. \tag{2}$$

For, let D be a cubical division of norm d. Let $d_1, d_2, \cdots$ be the cells of D containing points of $\mathfrak{A}$; while M_i, m_i denote as usual the maximum and minimum of f in d_i. Let $|f| \leqq F$ in $\mathfrak{A}$. For convenience, we also take $F \leqq 1$. Then for each $\varepsilon > 0$ there exists a d such that

$$\left| \bar{S}_D - \int_{\mathfrak{A}}^{\bar{}} \right| < \frac{\epsilon}{2}. \tag{3}$$

We may also take d so small that

$$\bar{\mathfrak{A}}_D - \bar{\mathfrak{A}} < \frac{\epsilon}{8F}. \tag{4}$$

Consider now the division Δ. Those of its cells containing points of $\mathfrak{A}$ fall into two classes: 1° those lying in only one cell of D; 2° those lying in more than one cell of D. Let $\delta_{i1}, \delta_{i2}, \cdots$ be the cells of the 1° class lying in d_i; let $\delta'_1, \delta'_2, \cdots$ denote all the cells of the 2° class.
Then the content of all the cells of Δ containing points of $\mathfrak{A}$ is

$$\bar{\mathfrak{A}}_\Delta = \sum \delta_{ik} + \sum \delta'_i. \tag{5}$$

But since the frontier of $\bar{\mathfrak{A}}_D$ is discrete, there exists a δ_0 such that

$$\sum \delta'_i < \frac{\epsilon}{16F}, \tag{6}$$

for any $\delta \leqq \delta_0$. As moreover $\bar{\mathfrak{A}}_\Delta$ converges to $\bar{\mathfrak{A}}$, we may also suppose δ_0 small enough that

$$\bar{\mathfrak{A}}_\Delta - \bar{\mathfrak{A}} < \frac{\epsilon}{16F}, \qquad \delta \leqq \delta_0. \tag{7}$$

From (5), (6) and (7) we have

$$|\sum \delta_{ik} - \bar{\mathfrak{A}}| < \frac{\epsilon}{8F},$$

which with (4) gives finally

$$|\sum \delta_{ik} - \bar{\mathfrak{A}}_D| < \frac{\epsilon}{4F}. \tag{8}$$

Let $\mathfrak{M}_{ik}$, $\mathfrak{M}'_i$ be the maxima of f in δ_{ik}, δ'_i respectively. Then

$$\mathfrak{M}_{ik} \leqq M_i, \qquad \mathfrak{M}_i \leqq F.$$

Hence

$$\bar{S}_\Delta \leqq \sum M_i \delta_{ik} + F \sum \delta'_i .$$

With (6) this gives

$$| \bar{S}_\Delta - \sum M_i \delta_{ik} | < \frac{\epsilon}{4} . \tag{9}$$

But

$$| \bar{S}_D - \sum M_i \delta_{ik} | \leqq F(\sum d_i - \sum \delta_{ik}) = F(\mathfrak{A}_D - \sum \delta_{ik})$$

$$< \frac{\epsilon}{4} , \text{ by (8).} \tag{10}$$

From (3), (9) and (10) we have

$$| \bar{S}_\Delta - \bar{\int_{\mathfrak{A}}} | < \epsilon , \qquad \delta \leqq \delta_0 .$$

This proves (1); in a similar manner we may establish (2).

It is now easy to prove the following theorems:

THEOREM 13. *Let $f(x_1, \cdots, x_m)$ be defined over the measurable field $\mathfrak{A}$. Let Δ be an unmixed division of $\mathfrak{A}$ of norm δ. Let $\underline{S}$ be the maximum of $\underline{S}_\Delta$ and $\bar{S}$ be the minimum of $\bar{S}_\Delta$ for all divisions Δ of norm $\delta \leqq \delta_0$. Then*

$$\underline{S} = \underline{\int_{\mathfrak{A}}} f d\mathfrak{A} , \qquad \bar{S} = \bar{\int_{\mathfrak{A}}} f d\mathfrak{A} .$$

THEOREM 14. *Let $f(x_1, \cdots, x_m)$ be integrable in $\mathfrak{A}$. Let $\mathfrak{B}_\eta$ be a partial aggregate depending on $\eta > 0$ such that $\bar{\mathfrak{B}}_\eta \doteq \bar{\mathfrak{A}}$ as $\eta \doteq 0$. Then*

$$\lim_{\eta = 0} \int_{\mathfrak{B}} f d\mathfrak{B} = \int_{\mathfrak{A}} f d\mathfrak{A} .$$

THEOREM 15. *Let $f(x_1, \cdots, x_m)$ be defined over $\mathfrak{A}$. Let $\mathfrak{B}$ be a partial aggregate of $\mathfrak{A}$ such that $\bar{\mathfrak{A}} = \bar{\mathfrak{B}}$. Then*

$$\bar{\int_{\mathfrak{A}}} f d\mathfrak{A} = \bar{\int_{\mathfrak{B}}} f d\mathfrak{B} ; \qquad \underline{\int_{\mathfrak{A}}} f d\mathfrak{A} = \underline{\int_{\mathfrak{B}}} f d\mathfrak{B} .$$

THEOREM 16. *Let $f(x_1, \cdots, x_m)$ be defined over the measurable field $\mathfrak{A}$, which is an inner field of $\mathfrak{B}$. Let $g(x_1, \cdots, x_m) = 0$ in $\mathfrak{B}$, except at the points of $\mathfrak{A}$ where it $= f(x_1, \cdots, x_m)$. Then*

$$\bar{\int_{\mathfrak{A}}} f d\mathfrak{A} = \bar{\int_{\mathfrak{B}}} g d\mathfrak{B} ; \qquad \underline{\int_{\mathfrak{A}}} f d\mathfrak{A} = \underline{\int_{\mathfrak{B}}} g d\mathfrak{B} .$$

THEOREM 17. *Let $f(x_1, \cdots, x_m)$ be defined over the field $\mathfrak{A}$ which is divided into the unmixed fields $\mathfrak{A}_1, \cdots, \mathfrak{A}_s$. Then*

$$\overline{\int_{\mathfrak{A}}} f d\mathfrak{A} = \overline{\int_{\mathfrak{A}_1}} + \cdots + \overline{\int_{\mathfrak{A}_s}},$$

$$\underline{\int_{\mathfrak{A}}} f d\mathfrak{A} = \underline{\int_{\mathfrak{A}_1}} + \cdots + \underline{\int_{\mathfrak{A}_s}}.$$

THEOREM 18. *Let $f(x_1, \cdots, x_m) \geqq \lambda$ in $\mathfrak{A}$. Then*

$$\overline{\int_{\mathfrak{A}}} f d\mathfrak{A} \geqq \lambda \bar{\mathfrak{A}}.$$

THEOREM 19. *Let $f(x_1, \cdots, x_m)$ be defined over the field $\mathfrak{A}$. Let $\mathfrak{A}_\sigma$ denote the points of $\mathfrak{A}$ at which $|f| \geqq \sigma$. If $\mathfrak{A}_\sigma$ is discrete for any $\sigma > 0$, f is integrable and*

$$\int_{\mathfrak{A}} f d\mathfrak{A} = 0.$$

THEOREM 20. *Let $f(x_1, \cdots, x_m)$ be integrable in the complete field $\mathfrak{A}$. Let $\mathfrak{C}$ be the points of $\mathfrak{A}$ at which f is continuous. Then $\bar{\mathfrak{C}} = \bar{\mathfrak{A}}$.*

For, if $\mathfrak{A}$ is discrete, the theorem is true even if f has no points of continuity in $\mathfrak{A}_i$. Let us therefore suppose that $\bar{\mathfrak{A}} > 0$. Let $\mathfrak{B}$ be the sifted aggregate of $\mathfrak{A}$. Let D be a rectangular division, and d one of its cells containing points of $\mathfrak{B}$. We can choose D so that no cell has points of $\mathfrak{B}$ only on its sides. Let $\mathfrak{a}$ be the points of $\mathfrak{A}$ in d. Since f is integrable in $\mathfrak{A}$, it is in $\mathfrak{a}$. We can therefore choose a system of cubical subdivisions of $\mathfrak{a}$ whose norms converge to zero, such that for each subdivision there is a cube lying in one of the preceding division, in which the oscillation is as small as we choose. These cubes converge to a point in $\mathfrak{a}$ at which f is continuous. Returning now to the original division D, we see that every cell of D which contains a point of $\mathfrak{B}$ contains a point of $\mathfrak{C}$. Hence $\bar{\mathfrak{C}}_D = \bar{\mathfrak{B}}_D$; hence $\bar{\mathfrak{C}} = \bar{\mathfrak{B}} = \bar{\mathfrak{A}}$.

THEOREM 21. *Let $f(x_1, \cdots, x_m) \geqq 0$ in $\mathfrak{A}$. If*

$$\int_{\mathfrak{A}} f d\mathfrak{A} = 0,$$

the points $\mathfrak{A}_\sigma$ at which f is equal to or greater than an arbitrarily small positive number σ, form a discrete aggregate. Let $\mathfrak{Z}$ denote the points at which $f = 0$. If $\mathfrak{A}$ is complete, $\bar{\mathfrak{A}} = \bar{\mathfrak{Z}}$.

For,

$$0 \geqq \overline{\int_{\mathfrak{A}_\sigma}} \geqq \sigma \bar{\mathfrak{A}}_\sigma.$$

Hence, $\bar{\mathfrak{A}}_\sigma = 0$. To prove the second part of the theorem, let b be a point of the sifted aggregate $\mathfrak{B}$ of $\mathfrak{A}$, at which f is continuous. Then we can choose $\delta > 0$ so small that $f \equiv \lambda > 0$ in any $D_\delta(b)$. But the upper content of the points $\bar{a}$ of $\mathfrak{A}$ in $D(b)$ is $\bar{a} > 0$. Hence

$$\int_{\mathfrak{A}} f d\mathfrak{A} \equiv \int_{a} f da \equiv \lambda \bar{a} > 0.$$

Hence $f = 0$ at every point of continuity of $\mathfrak{A}$ in $\mathfrak{B}$. Let now D be a rectangular division of space. Any cell of D which contains a point of $\mathfrak{B}$ also contains a point of continuity lying in $\mathfrak{B}$. Hence

$$\bar{\mathfrak{Z}}_D \equiv \mathfrak{B}_D;$$

which gives

$$\bar{\mathfrak{Z}} \equiv \bar{\mathfrak{B}}, \qquad \text{or} \qquad \bar{\mathfrak{Z}} = \bar{\mathfrak{B}}.$$

§ 6. *Reduction of multiple integrals to iterated integrals.*

Let $\mathfrak{A}$ be a point aggregate in R_m. As $x = (x_1, \cdots, x_m)$ ranges over $\mathfrak{A}$, x_i will range over an aggregate $\underline{\mathfrak{x}}_i$ on the x_i axis which we call the *projection of* $\mathfrak{A}$ *on this axis.*

The points of R_m for which one of the coördinates as x_i has a fixed value $x_i = \xi_i$ lie in an $(m-1)$-way plane perpendicular to the x_i axis. We denote it by P_{ξ_i} or more shortly by P_i. The points of $\mathfrak{A}$ in P_{ξ_i} form a *plane section of* $\mathfrak{A}$ *corresponding to the point* ξ_i in $\underline{\mathfrak{x}}_i$. We denote it by $\mathfrak{P}_{\xi_i}$ or by $\mathfrak{P}_i$.

As $x = (x_1, \cdots, x_m)$ ranges over $\mathfrak{A}$, the point $(x_1, \cdots, x_{i-1}, 0, x_{i+1}, \cdots, x_m)$ ranges over an aggregate $\mathfrak{X}_i$, called *the projection of* $\mathfrak{A}$ *on the plane* $x_i = 0$.

Let us fix all the coördinates of $x = (x_1, \cdots, x_m)$ except x_i. Then x describes a line; the points of $\mathfrak{A}$ on one of these lines we call the *rectilinear section of* $\mathfrak{A}$ *parallel to the* x_i *axis.* We mention a few properties of these aggregates, whose demonstration is readily given.

THEOREM 22. *If* $\mathfrak{A}$ *is complete, so are the aggregates* $\underline{\mathfrak{x}}_i$, $\mathfrak{X}_i$, $\mathfrak{a}_i$, $\mathfrak{P}_i$, *above mentioned.*

THEOREM 23. *Let* $\mathfrak{A}$ *be measurable. Let* $\underline{\mathfrak{x}}_\sigma$ *denote those points of* $\underline{\mathfrak{x}}_i$ *for which the upper content of the frontier points of* $\mathfrak{P}_i$ *is* $\equiv \sigma$. *Then* $\underline{\mathfrak{x}}_\sigma$ *is discrete.*

A similar theorem holds for the aggregate $\mathfrak{X}_\sigma$.

THEOREM 24. *Let* $f(x_1, \cdots, x_m)$ *be defined over the measurable field* $\mathfrak{A}$. *Then*

$$(1) \qquad \underline{\int_{\mathfrak{A}}} f d\mathfrak{A} \equiv \underline{\int_{\mathfrak{X}_i}} dx_i \underline{\int_{\mathfrak{P}_i}} f d\mathfrak{P}_i \equiv \overline{\int_{\mathfrak{X}_i}} dx_i \underline{\int_{\mathfrak{P}_i}} f d\mathfrak{P}_i \equiv \overline{\int_{\mathfrak{B}}} f d\mathfrak{A},$$

$$(2) \qquad \underline{\int_{\mathfrak{A}}} f d\mathfrak{A} \equiv \underline{\int_{\mathfrak{X}_i}} dx_i \overline{\int_{\mathfrak{P}_i}} f d\mathfrak{P} \equiv \overline{\int_{\mathfrak{X}_i}} dx_i \overline{\int_{\mathfrak{P}_i}} f d\mathfrak{P}_i \equiv \overline{\int_{\mathfrak{A}}} f d\mathfrak{A}.$$

Let us prove (1); the demonstration of (2) is similar. Let $\mathfrak{A}$ lie in an outer cube $\mathfrak{B}$, whose projection on the x_i axis is $\mathfrak{b}$, and whose plane sections perpendicular to this axis are $\mathfrak{Q}$. We introduce the auxiliary function

$$g(x_1, \cdots, x_m) = f(x_1, \cdots, x_m) \text{ in } \mathfrak{A},$$
$$= 0 \text{ at other points of } \mathfrak{B}.$$

Let D be a cubical division of R_m of norm d. This divides $\mathfrak{B}$ into cells which we denote by δ. It also divides the plane section $\mathfrak{Q}$ into cells δ', and the segment $\mathfrak{b}$ into intervals which we denote by δ''.

Let M, M' be the maxima, and m, m' the minima of g in the cells δ, δ'. Let $|f| \leqq F$ in $\mathfrak{A}$. Let $\bar{G}$, $\underline{G}$ denote the upper and lower integrals of g in the field $\mathfrak{B}$. Then for each $\epsilon > 0$ there exists a d_0 such that

$$\underline{G} - \epsilon < \sum_{\mathfrak{B}} m\delta \leqq \sum_{\mathfrak{B}} M\delta < \bar{G} + \epsilon \qquad (d \leq d_0). \tag{3}$$

Also for each x_i of $\mathfrak{b}$

$$\sum_{\mathfrak{Q}} m'\delta' \leqq \underline{\int}_{\mathfrak{Q}} g d\mathfrak{Q} \leqq \sum_{\mathfrak{Q}} M'\delta';$$

or since $m \leqq m'$, and $M' \leqq M$,

$$\sum_{\mathfrak{Q}} m\delta' \leqq \underline{\int}_{\mathfrak{Q}} g d\mathfrak{Q} \leqq \sum_{\mathfrak{Q}} M\delta'.$$

Multiplying by δ'' and summing over $\mathfrak{b}$ we have, since $\delta = \delta'\delta''$,

$$\sum_{\mathfrak{B}} m\delta \leqq \sum_{\mathfrak{b}} \delta'' \underline{\int}_{\mathfrak{Q}} \leqq \sum_{\mathfrak{B}} M\delta.$$

Combined with (3), this gives

$$\underline{G} - \epsilon < \sum_{\mathfrak{b}} \delta'' \underline{\int}_{\mathfrak{Q}} < \bar{G} + \epsilon,$$

or

$$\underline{G} \leqq \underline{\int}_{\mathfrak{b}} \underline{\int}_{\mathfrak{Q}} \leqq \int_{\mathfrak{b}} \underline{\int}_{\mathfrak{Q}} \leqq \bar{G}. \tag{4}$$

From (4) and theorem 16 we have

$$\underline{\int}_{\mathfrak{A}} f d\mathfrak{A} \leqq \underline{\int}_{\mathfrak{b}} dx_i \int_{\mathfrak{Q}} g d\mathfrak{Q} \leqq \bar{\int}_{\mathfrak{b}} dx_i \underline{\int}_{\mathfrak{Q}} g d\mathfrak{Q} \leqq \bar{\int}_{\mathfrak{A}} f d\mathfrak{A}. \tag{5}$$

Let $\bar{\mathfrak{F}}_{\mathfrak{P}}$ denote the upper content of the frontier points of $\mathfrak{P}_i$. Then

$$\left| \underline{\int}_{\mathfrak{P}_i} f d\mathfrak{P} - \underline{\int}_{\mathfrak{Q}} g d\mathfrak{Q} \right| \leqq F \bar{\mathfrak{F}}_{\mathfrak{P}}. \tag{6}$$

Let us set

$$h(x_i) = \underline{\int}_{\mathfrak{P}_i} f d\mathfrak{P}_i, \text{ for points of } \mathfrak{x}_i;$$

$$= 0, \text{ for other points of } \mathfrak{b}.$$

Then by (6), since the points of $\mathfrak{x}_i$ for which $\bar{\mathfrak{F}}_{\mathfrak{P}} > \sigma$ form a discrete set,

$$\underline{\int}_{\mathfrak{b}} h dx_i = \underline{\int}_{\mathfrak{x}} dx_i \underline{\int}_{\mathfrak{Q}} g d\mathfrak{Q}. \tag{7}$$

But obviously

$$\underline{\int}_{\mathfrak{b}} h dx_i = \underline{\int}_{\mathfrak{x}_i} h dx_i. \tag{8}$$

Hence from (7), (8),

$$\underline{\int}_{\mathfrak{x}} dx_i \underline{\int}_{\mathfrak{Q}} g d\mathfrak{Q} = \underline{\int}_{\mathfrak{x}_i} dx_i \underline{\int}_{\mathfrak{P}_i} f d\mathfrak{P}_i. \tag{9}$$

In the same way we show

$$\bar{\int}_{\mathfrak{x}} dx_i \underline{\int}_{\mathfrak{Q}} g d\mathfrak{Q} = \bar{\int}_{\mathfrak{x}_i} dx_i \underline{\int}_{\mathfrak{P}} f d\mathfrak{P}_i. \tag{10}$$

By placing (9) and (10) in (5) we get (1).

In a similar manner we may prove:

THEOREM 25. *Let $\mathfrak{A}$ be measurable. Let $\mathfrak{X}_i$ be its projection on the plane $x_i = 0$. Let $\mathfrak{a}_i$ be the rectilinear section of $\mathfrak{A}$ parallel to the x_i axis. Then*

$$\underline{\int}_{\mathfrak{A}} f d\mathfrak{A} \leqq \underline{\int}_{\mathfrak{X}_i} d\mathfrak{X}_i \underline{\int}_{\mathfrak{a}_i} f dx_i \leqq \bar{\int}_{\mathfrak{X}_i} d\mathfrak{X}_i \underline{\int}_{\mathfrak{a}_i} f dx_i \leqq \bar{\int}_{\mathfrak{A}} f d\mathfrak{A}.$$

$$\underline{\int}_{\mathfrak{A}} f d\mathfrak{A} \leqq \underline{\int}_{\mathfrak{X}_i} d\mathfrak{X}_i \bar{\int}_{\mathfrak{a}_i} f dx_i \leqq \bar{\int}_{\mathfrak{X}_i} d\mathfrak{X}_i \bar{\int}_{\mathfrak{a}_i} f dx_i \leqq \bar{\int}_{\mathfrak{A}} f d\mathfrak{A}.$$

Many theorems follow now from 24 and 25 as corollaries. We state only the two following.

THEOREM 26. *Let $f(x_1, \cdots, x_m)$ be integrable in the measurable field $\mathfrak{A}$. Let $\mathfrak{x}_\sigma$ denote those points of $\mathfrak{x}_i$ for which*

$$\bar{\int}_{\mathfrak{P}_i} f d\mathfrak{P} - \underline{\int}_{\mathfrak{P}} f d\mathfrak{P} \geqq \sigma.$$

Then $\mathfrak{x}_\sigma$ is discrete. If $\mathfrak{A}$ is complete, the upper content of the points at which the above difference vanishes is $\bar{\mathfrak{x}}_i$.

THEOREM 27. *Let $f(x_1, \cdots, x_m)$ be integrable in the measurable complete field $\mathfrak{A}$. Let $\mathfrak{y}_i$ denote those points of $\mathfrak{x}_i$ for which the integrals over the cor-*

responding plane sections $\mathfrak{P}_i$ exist. Let $\mathfrak{Y}_i$ denote the points of $\mathfrak{X}_i$ for which the integrals over the corresponding rectilinear sections $\mathfrak{a}_i$ exist. Then

$$\int_{\mathfrak{A}} f d\mathfrak{A} = \int_{\mathfrak{y}_i} dx_i \int_{\mathfrak{P}_i} f d\mathfrak{P}_i = \int_{\mathfrak{y}_i} d\mathfrak{Y}_i \int f dx_i .$$

§ 7. *Change of variables.*

Let the transformation T be defined by the equations

$$T; \qquad x_1 = \phi_1(t_1, \cdots, t_m), \cdots, x_m = \phi_m(t_1, \cdots, t_m),$$

where the ϕ's have continuous first derivatives in the region R. Let the correspondence between R and its image R_T be uniform. Let the determinant of the transformation $J \neq 0$. We say the transformation T is *regular*.

It is easy to show now that to any *inner* aggregate $\mathfrak{A}$ in R corresponds an inner aggregate $\mathfrak{A}_T$ in R_T. To inner and frontier points of $\mathfrak{A}$ correspond similar points in $\mathfrak{A}_T$, and conversely.

As a direct consequence of theorem 6 we have

THEOREM 28. *If either $\mathfrak{A}$ or $\mathfrak{A}_T$ is measurable, so is the other. Also if one is discrete, so is the other.*

We show next how the transformation T can be expressed as the product of two simpler *regular* transformations.

Since $J \neq 0$, not all the first partial derivatives vanish at any point of R. Let

$$\frac{\partial \phi_m}{\partial t_m} \neq 0$$

at a point t and hence in a certain domain $D_\delta(t)$ of t. Let us define the first of these new transformations by

$$T_1; \qquad u_1 = t_1, \cdots, u_{m-1} = t_{m-1}, \qquad u_m = \phi_m(t_1, \cdots, t_m).$$

Since T_1 is regular, it can be inverted, giving

$$t_1 = u_1, \cdots, t_{m-1} = u_{m-1}, \; t_m = \theta(u_1, \cdots, u_m),$$

where θ is a one-valued function having continuous first partial derivatives in a certain domain $D_\eta(u)$. If $\delta' \leqq \delta$ is taken sufficiently small, the image U of $D_{\delta'}(t)$ lies in $D_\eta(u)$. The second transformation is

$$T_2; \quad x_1 = \phi_1(u_1, \cdots, u_{m-1}, \theta), \cdots, x_{m-1} = \phi_{m-1}(u_1, \cdots, u_{m-1}, \theta), \quad x_m = u_m,$$

and is defined over U.

Obviously now

$$T = T_1 T_2$$

when t ranges over $D_{\delta'}(t)$, and the determinant J_2 of T_2 is $\neq 0$, since the determinants J, J_1 of T and T_1 are $\neq 0$, by hypothesis.

Let $\mathfrak{A}$ be an inner aggregate of R. It is easy now to show that we can effect a cubical division of the t-space of norm d such that for the points of $\mathfrak{A}$ in each cell d_k, there exist two transformations $T_1^{(k)}$, $T_2^{(k)}$ of the kind just considered such that

$$T = T_1^{(k)} T_2^{(k)}. \tag{1}$$

For we can take d so small that not all the first partial derivatives vanish in any cell, since otherwise $J = 0$ at some point in this cell. Next, making use of the uniform continuity of our functions, we show that the norms η of the domains $D_\eta(u)$, and hence the norms δ' of $D_{\delta'}(t)$ considered above, do not sink below some positive number, as t ranges over any inner region of R. Thus taking $d > 0$ small enough, the relation (1) holds in each cell containing a point of $\mathfrak{A}$.

THEOREM 29. *Let*

$$T; \qquad x_1 = \phi_1(t_1, \cdots, t_m), \ \cdots, \ x_m = \phi_m(t_1, \cdots, t_m)$$

define a regular transformation of determinant J in the region R. Let $\mathfrak{T}$ be any inner perfect measurable aggregate of R and let $\mathfrak{X}$ be its image. Let $f(x_1, \cdots, x_m)$ be continuous in $\mathfrak{X}$. Then

$$\int_{\mathfrak{X}} f(x_1, \cdots, x_m)\, dx_1 \cdots dx_m = \int_{\mathfrak{T}} |J| f dt_1 \cdots dt_m. \tag{2}$$

For $m = 1$, the theorem is obviously true. Let us therefore assume it correct for $m - 1$, and show it is so for m. As just shown, we can effect a cubical division of the t-space such that the relation (1) holds in each cell containing points of $\mathfrak{T}$. But if (2) holds for each of these partial aggregates of $\mathfrak{T}$, it obviously holds for $\mathfrak{T}$. We may therefore assume without loss of generality that the relation (1) holds throughout $\mathfrak{T}$. Let $\mathfrak{U}$ be the image of $\mathfrak{X}$ in the u-space, while $\mathfrak{y}'_m$, $\mathfrak{P}'_m$ are the images of $\mathfrak{y}_m$, $\mathfrak{P}_m$. Then

$$\begin{aligned}
\int_{\mathfrak{X}} f dx_1 \cdots dx_m &= \int_{\mathfrak{y}_m} dx_m \int_{\mathfrak{P}_m} f dx_1 \cdots dx_{m-1} \\
&= \int_{\mathfrak{y}'_m} du_m \int_{\mathfrak{P}'_m} |J_2| f du_1 \cdots du_{m-1} \\
&= \int_{\mathfrak{U}} |J_2| \, du_1 \cdots du_m.
\end{aligned}$$

The application of the transformation T_2 to the integral on the right gives similarly,

$$\int_{\mathfrak{U}} |J_2| du_1 \cdots du_m = \int_{\mathfrak{T}} |J_1||J_2| f dt_1 \cdots dt_m .$$

This and the preceding equation give (2).

As a corollary of the last theorem we have

THEOREM 30. *Let d_k denote the volume of one of the cells of a rectangular division D of norm d in the t-space. Let $\mathfrak{T}$ be any inner region of R and $\mathfrak{X}$ its image. To the division D correspond a division Δ of the x-space into cells of volume δ_k. The cells of Δ following within $\mathfrak{X}$ are unmixed, and*

$$\delta_k = |J| d_k + \epsilon_k, \qquad t \text{ in } d_k,$$

where $|\epsilon_k| < \epsilon$ uniformly, if d is taken sufficiently small.

THEOREM 31. *Let T be a regular transformation of determinant J defined over a region R. Let $\mathfrak{T}$ be an inner aggregate of R and let $\mathfrak{X}$ be its image. Let $f(x_1, \cdots, x_m)$ be limited in $\mathfrak{T}$. Then*

$$\bar{\int}_{\mathfrak{X}} f(x_1, \cdots, x_m) d\mathfrak{X} = \bar{\int}_{\mathfrak{T}} |J| f d\mathfrak{T}, \tag{3}$$

$$\underline{\int}_{\mathfrak{X}} f(x_1, \cdots, x_m) d\mathfrak{X} = \underline{\int}_{\mathfrak{T}} |J| f d\mathfrak{T}. \tag{4}$$

Let us prove (3); the demonstration of (4) is similar. To the end we effect a cubical division of the t-space of norm d. To it corresponds a division of $\mathfrak{X}$ into cells of norm δ.

Let us consider the integral on the left of (3). Employing the usual notation, and letting J_i denote the value of J at some point of d_i, we have

$$\begin{aligned} \sum M_i \delta_i &= \sum M_i (|J_i| + \epsilon_i) d_i \\ &= \sum M_i |J_i| d_i + \sum \epsilon_i M_i d_i . \end{aligned} \tag{5}$$

Here

$$|\sum \epsilon_i M_i d_i| < \epsilon F \overline{\text{cont}}\ \mathfrak{T} = \eta, \tag{6}$$

which may be made as small as we choose, on taking d sufficiently small.

Let us consider now the integral on the right of (3). In the cell d_i,

$$\max f \cdot \min |J| \leqq \max \cdot f|J| \leqq \max f \cdot \max |J| ,$$

if $\max f$ is positive, while the signs are reversed if it is negative. Let us set

$$M_i' = \max \cdot f|J|, \text{ in } d_i .$$

Then

$$M_i' = M_i (|J_i| + \epsilon_i'),$$

where $|\epsilon_i'| < \epsilon$ uniformly.

Thus

$$\sum M_i' d_i = \sum M_i |J| d_i + \sum \epsilon_i' M_i d_i, \tag{7}$$

where, as in (6),

$$|\sum \epsilon_i' M_i d_i| < \eta.$$

Thus (5) and (7) give

$$|\sum M_i \delta_i - \sum M_i d_i| < 2\eta,$$

which proves (3).

OUTER ISLAND, *May*, 1905.